

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES  
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum  
Internationales Büro



(43) Internationales Veröffentlichungsdatum  
25. September 2003 (25.09.2003)

PCT

(10) Internationale Veröffentlichungsnummer  
WO 03/078629 A1

(51) Internationale Patentklassifikation<sup>7</sup>: C12N 15/11, 15/82

(21) Internationales Aktenzeichen: PCT/EP03/02735

(22) Internationales Anmeldedatum:  
17. März 2003 (17.03.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:  
102 12 892.8 20. März 2002 (20.03.2002) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von  
US): BASF PLANT SCIENCE GMBH [DE/DE]; Carl-  
Bosch-Strasse 38, 67056 Ludwigshafen (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (nur für US): KOCK, Michael  
[DE/DE]; Am Leutbusch 12, 67105 Schifferstadt (DE).  
BAUER, Jörg [DE/DE]; Friedrich-Profit-Str. 56, 67063  
Ludwigshafen (DE).

(74) Anwalt: DÖRPER, Thomas; BASF Aktiengesellschaft,  
67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten (national): AE, AG, AL, AM, AT,  
AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR,  
CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE,  
GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,  
KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK,  
MN, MW, MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO,  
RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ,  
UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (regional): ARIPO-Patent (GH,  
GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),  
eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ,  
TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE,  
DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL,  
PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG,  
CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

- mit internationalem Recherchenbericht
- vor Ablauf der für Änderungen der Ansprüche geltenden  
Frist; Veröffentlichung wird wiederholt, falls Änderungen  
eintreffen

Zur Erklärung der Zweibuchstaben-Codes und der anderen  
Abkürzungen wird auf die Erklärungen ("Guidance Notes on  
Codes and Abbreviations") am Anfang jeder regulären Ausgabe  
der PCT-Gazette verwiesen.

(54) Title: CONSTRUCTS AND METHODS FOR THE REGULATION OF GENE EXPRESSION

(54) Bezeichnung: KONSTRUKTE UND VERFAHREN ZUR REGULATION DER GENEXPRESSION

(57) Abstract: The invention relates to constructs and methods for the regulation of gene expression of at least two endogenous target genes by introduction of an at least partly double-stranded ribonucleic acid molecule into a eukaryotic cell or a eukaryotic organism, whereby the ribonucleic acid molecule comprises at least two ribonucleotide sequence sections which are homologous with various genes of the eukaryotic cell.

(57) Zusammenfassung: Die vorliegende Erfindung betrifft Konstrukte und Verfahren zur Regulation der Genexpression von mindestens zwei endogenen Zielgenen durch Einbringen eines zumindest teilweise doppelsträngigen Ribonukleinsäuremoleküls in eine eukaryotische Zelle oder einen eukaryotischen Organismus, wobei das Ribonukleinsäuremolekül mindestens zwei Ribonukleotidsequenzabschnitte umfasst, die zu verschiedenen Genen der eukaryotischen Zelle homolog sind.

WO 03/078629 A1

## Konstrukte und Verfahren zur Regulation der Genexpression

## Beschreibung

5.

- Die vorliegende Erfindung betrifft Konstrukte und Verfahren zur Regulation der Genexpression von mindestens zwei endogenen Zielgenen durch Einbringen eines zumindest teilweise doppelsträngigen Ribonukleinsäuremoleküls in eine eukaryotische Zelle oder einen eukaryotischen Organismus, wobei das Ribonukleinsäuremolekül mindestens zwei Ribonukleotidsequenzabschnitte umfasst, die zu verschiedenen Genen der eukaryotischen Zelle homolog sind.

- Die gezielte Inhibition der Genexpression definierter Gene ist eine der am meisten beforschten Technologie der Biotechnologie. Die Expression von antisense-RNA ist dabei der am häufigsten verwendete Ansatz und vielfach beschrieben (u.a. EP-A1 0 458 367; EP-A1 0 140 308; van der Krol AR et al. (1988) BioTechniques 6(10):658-676; de Lange P et al. (1995) Curr Top Microbiol Immunol 197:57-75). Antisense-RNA vermittelte Ansätze haben jedoch den Nachteil, dass stöchiometrische Mengen der antisense-RNA erforderlich sind, um eine wirksame Inhibition der Ziel-mRNA zu bewirken. Weitere Probleme stehen im Zusammenhang mit dem Einbringen der antisense-RNA in ausreichenden Mengen in die Zellen und mit der Labilität der antisense-RNA. Ansätze basierend auf antisense-RNA sind daher meist ineffizient.

- Ein weiterer Ansatz zur Genregulation ist die "Co-Suppression" und meint die Verminderung der Expression eines endogenen Zielgens durch transgene Expression einer sense-RNA dieses Zielgens (EP-A1 0 465 572). Der Co-Suppression liegen vermutlich mehr als ein Mechanismus zugrunde. Nachteilig ist die mangelnde Verlässlichkeit und Reproduzierbarkeit des Verfahrens. In manchen Fällen erfolgt Suppression, während in anderen Fällen - bedingt durch die Expression der sense-RNA - die erwartete Überexpression erfolgt. Auch ist der erhaltene Phänotyp oft nicht stabil. Die Anwendung der Co-Suppression ist im wesentlichen auf Pflanzen beschränkt.

- Verschiedene Abwandlungen der Verfahren basierend auf antisense-RNA oder Cosuppression sind bekannt. So beschreibt WO 93/23551 ein Verfahren zur Inhibition mehrerer Gene durch Expression einer chimären antisense-RNA oder sense-RNA. Das Verfahren kann die üblichen mit antisense-RNA oder sense-RNA verbundenen Probleme nicht lösen und bleibt ineffizient.

WO 98/36083 und WO 99/15682 beschreiben die Regulation der Genexpression mittels viraler Expressionssysteme ("virus induced gene

silencing" VIGS).

WO 99/32619 und WO 99/53050 beschreiben Verfahren zur Inhibition einzelner Zielgene unter Verwendung einer RNA mit doppelsträngiger Struktur, wobei das Zielgen und die Region der RNA Duplex zu-  
5     mindest eine teilweise Identität aufweisen (siehe auch: Montgomery MK et al. (1998) Proc Natl Acad Sci USA 95:15502- 15507; Sharp PA (1999) Genes & Development 13(2):139-141; Fire A et al. (1998) Nature 391:806-11). Das Verfahren wird heute auch als  
10    "RNA-Interference" (RNAi) bezeichnet und hat in Mechanismus und Wirkung Ähnlichkeiten mit dem oben erwähnten VIGS Verfahren.

Die beschriebenen Verfahren, insbesondere das RNAi-Verfahren, lösen zwar einige Probleme im Zusammenhang mit der Verminderung  
15    einzelner Zielgene. Für andere Probleme, insbesondere für die parallele Suppression mehrerer Zielgene, konnte jedoch bislang keine befriedigende Lösung bereit gestellt werden. Zahlreiche Ansätze in der Biotechnologie erfordern nicht nur die Verminderung eines einzelnen Zielgens, sondern mehrerer Zielgene, wie bei-  
20    spielsweise verschiedener Gene eines oder verschiedener Stoffwechselwege oder ganzer Genfamilien. Bislang war dies nur mit erheblichen Arbeits- und Zeitaufwand zu realisieren. Die Ansätze erforderten oft die individuelle Regulation der einzelnen Zielgene durch sukzessive Transformation beispielsweise mit verschiedenen Expressionskonstrukten, die jeweils für eine antisense RNA  
25    eines Zielgens kodierte. Neben dem erheblichen Arbeits- und Zeitaufwand, besteht dabei der Nachteil, das für viele Systeme und Organismen nur eine beschränkte Anzahl von Selektionsmarkern, geeigneten Promotoren etc. zur Verfügung steht, was multiple Transformationen erheblich erschwert und beispielsweise die Deletion der Marker nach der Transformation und Selektion erfordert.  
30    Die mehrfache Verwendung eines Promotors hat oft unerwünschte Folgen, wie beispielsweise ein epigenetisches Gene-Silencing. Hierbei kommt es infolge der mehrfach verwendeten Kontrollsequenzen zu einer Inaktivierung derselben, vergleichbar der oben be-  
35    schriebenen Cosuppression.

Es stellte sich also die Aufgabe, neue Verfahren bereit zu stellen, die eine effiziente Verminderung der Expression mindestens zweier endogener Zielgene in einer eukaryotischen Zelle oder ei-  
40    nem eukaryotischen Organismus ermöglichen. Diese Aufgabe wird durch die vorliegende Erfindung gelöst.

Ein erster Gegenstand der Erfindung betrifft Verfahren zur Verminderung der Expression von mindestens zwei verschiedenen, endo-  
45    genen Zielgenen in einer eukaryotischen Zelle oder einem eukaryotischen Organismus durch Einbringen eines zumindest teilweise doppelsträngigen Ribonukleinsäuremoleküls in besagte eukaryotische

## 3

Zelle oder besagten eukaryotischen Organismus, wobei das doppelsträngige Ribonukleinsäuremolekül umfasst

- 5 a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei jeweils mindestens eine dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines jeden der besagten endogenen Zielgene und
- 10 b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.

Ein weiterer Gegenstand der Erfindung umfasst ein zumindest teilweise doppelsträngiges Ribonukleinsäuremolekül, wobei das doppelsträngige Ribonukleinsäuremolekül umfasst

- 20 a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei jeweils mindestens eine dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens, wobei jedoch nicht alle "sense"-Ribonukleotidsequenzen zu dem "sense"-RNA-Transkript eines einzigen endogenen Zielgens identisch sind, und
- 25 b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.

30 Umfasst ist ferner die Verwendung der erfindungsgemäßen doppelsträngiges Ribonukleinsäuremolekül in einem der erfindungsgemäßen Verfahren.

Die vorliegende Erfindung löst die oben geschilderten Probleme und ermöglicht eine schnelle, besonders effiziente Methode zur Regulation der Expression verschiedener Zielgene. Insbesondere ergeben sich folgende Vorteile:

- 40 a) Transgene Organismen oder Zellen, in denen mehr als ein Zielgen inhibiert wird, können in einer einzigen Transformation erzeugt werden.
- 45 b) Die Transkriptionsrate für jeden Ribonukleotidsequenz der dsRNA ist gleich. Dadurch werden multiple Phänotypen durch unterschiedliche Expressionshöhen verhindert, wie sie bei individueller Expression separater Ribonukleotidsequenzen - beispielsweise durch den unterschiedlichen Ort der Insertion



in das Genom - oft entstehen. Dieser Vorteil gewährleistet eine gleichbleibend hohe Inhibition aller Zielgene und vermindert dramatisch die erforderlichen Selektionsschritte zu Generierung eines Organismus, bei dem alle Zielgene effizient  
5 supprimiert werden.

- c) Ein ökonomischer Umgang mit Kontrollelementen wie Promotoren und Selektionsmarkern wird ermöglicht. Zudem erübrigen sich Probleme, wie sie bei der mehrfachen Verwendung eines bestimmten Kontrollelementes, insbesondere eines Promoters,  
10 entstehen können ("epigenic gene silencing").
- d) Eine Segregation der einzelnen Ribonukleotidsequenzen bei nachfolgenden Züchtungs- und Kreuzungsschritten, wie sie bei der Verwendung mehrerer Expressionskonstrukte zwangsläufig  
15 entsteht, wird verhindert. Dadurch wird die nachfolgende Züchtung stabiler Linien erheblich erleichtert und beschleunigt.
- 20 e) Organismen mit komplexen beispielsweise polyploide Genomen, wie beispielsweise manche Pflanzen, sind einer effizienten Gensuppression zugänglich. Aufgrund der zahlreichen Kopien für einzelne Gene sind diese Organismen klassischen verfahren der Mutagenese und Selektion nicht zugänglich.
- 25 Überraschenderweise konnte bei dem erfindungsgemäßen Verfahren keine störende Interferenz zwischen den einzelnen Ribonukleotidsequenzabschnitte untereinander beobachtet werden.
- 30 "Endogenes Zielgen einer eukaryotischen Zelle oder eines eukaryotischen Organismus" meint jede Nukleinsäuresequenz in einer eukaryotischen Zelle, einem eukaryotischen Organismus oder einem Teil, Organ, Gewebe, Samen etc. desselben, die zur Transkription befähigt ist. Dabei kann es sich um natürlicherweise vorkommende oder  
35 aber künstlich eingeführte Sequenzen (wie beispielsweise transgene Sequenzen) handeln, wobei natürlicherweise vorkommende Sequenzen bevorzugt sind. Natürlicherweise vorkommende Sequenzen sind bevorzugt und umfassen sowohl die eigenen Sequenzen der eukaryotischen Zelle oder des eukaryotischen Organismus als auch  
40 Gene von Pathogenen, die in der eukaryotischen Zelle oder dem eukaryotischen Organismus nach einem Befall durch ein Pathogen präsent sind. Das Zielgen kann in der chromosomalen DNA oder der DNA der Organellen (wie beispielsweise der Plastiden z.B. Chloroplasten etc.) lokalisiert sein oder aber sich extrachromosomal in  
45 der Zelle befinden. Die natürlicherweise vorkommenden, eigenen Sequenzen des eukaryotischen Organismus umfassen bevorzugt Gene desselben, die stabil im Genom vorliegen, wobei das Genom die Ge-

## 5

samtheit der genetischen Information meint und sowohl die chromosomale als auch die plastidäre DNA umfasst. Bevorzugt ist das endogene Zielgen ein natürlicherweise in der chromosomalen DNA vorkommendes Gen. Bevorzugt sind Gene deren verminderte Expression zu einem veränderten Phänotyp führt.

"Verminderung" oder "vermindern" der Expression eines Zielgens ist im Zusammenhang weit auszulegen und umfasst die teilweise oder im wesentlichen vollständige, auf unterschiedliche zellbiologische Mechanismen beruhende Unterbindung oder Blockierung der Expression des Zielgens oder der von ihm abgeleiteten RNA, mRNA, rRNA, tRNA und/oder des dadurch kodierten Proteinproduktes in einer Zelle oder einem Organismus oder einem davon abgeleiteten Teil, Gewebe, Organ, Zelle oder Samen. Eine Verminderung im Sinne der Erfindung umfasst die mengenmässige Verringerung einer vom Zielgen exprimierten RNA, mRNA, rRNA, tRNA und/oder des dadurch kodierten Proteinproduktes bis hin zu einem im wesentlichen vollständigen Fehlen derselben. Dabei wird die Expression einer bestimmten RNA, mRNA, rRNA, tRNA und/oder des dadurch kodierten Proteinproduktes in einer Zelle oder einem Organismus im Vergleich zu der selben Zelle oder Organismus, die dem Verfahren nicht unterworfen wurden, bevorzugt um mehr als 50%, besonders bevorzugt um mehr als 80%, ganz besonders bevorzugt um mehr als 90%, am meisten bevorzugt mehr als 95% vermindert. Dabei kann die Verminderung durch dem Fachmann geläufigen Verfahren ermittelt werden. So kann die Verminderung der Proteinmenge beispielsweise durch immunologischen Nachweis des Proteins bestimmt werden. Weiterhin können biochemische Techniken wie Northern-Hybridisierung, "nuclease protection assay", Reverse Transkription (quantitative RT-PCR), ELISA ("enzyme linked immunosorbent assay"), Western-Blotting, Radioimmunoassay (RIA) oder andere Immunoassays sowie "fluorescence activated cell analysis" (FACS) eingesetzt werden. Je nach Art des verminderten Proteinproduktes kann auch dessen Aktivität oder der Einfluss auf den Phänotyp des Organismus oder der Zelle ermittelt werden.

"Proteinmenge" meint die Menge eines bestimmten Polypeptides in einem Organismus, einem Gewebe, einer Zelle oder einem Zellkompartiment.

40

"Verminderung" der Proteinmenge meint die mengenmässige Verminderung der Menge eines bestimmten Polypeptides in einem Organismus, einem Gewebe, einer Zelle oder einem Zellkompartiment - beispielsweise durch das erfindungsgemässe Verfahren - im Vergleich zu dem Wildtyp derselben Gattung und Art auf den dieses Verfahren nicht angewendet wurde, unter ansonst gleichen Rahmenbedingungen (wie beispielsweise Kulturbedingungen, Alter, Nährstoffzufuhr

etc.). Der Verminderung beträgt dabei mindestens 10 %, bevorzugt mindestens 10% oder mindestens 20%, besonders bevorzugt um mindestens 40% oder 60%, ganz besonders bevorzugt um mindestens 70% oder 80%, am meisten bevorzugt um mindestens 90% oder 95%. Verfahren zur Bestimmung der Proteinmenge sind dem Fachmann bekannt. Beispielfhaft seien zu nennen: Das Mikro-Biuret Verfahren (Goa J (1953) Scand J Clin Lab Invest 5:218-222), die Folin-Ciocalteu-Methode (Lowry OH et al. (1951) J Biol Chem 193:265-275) oder die Messung der Adsorption von CBB G-250 (Bradford MM (1976) Analyt Biochem 72:248-254).

"Verschieden" meint in Bezug auf zwei endogene Zielgene bevorzugt, dass die von den beiden endogenen Zielgenen transkribierte RNA oder mRNA nicht identisch ist. Bevorzugt ist die Homologie der von den beiden endogenen Zielgenen transkribierte RNA oder mRNA geringer als 90%, bevorzugt geringer als 80%, besonders bevorzugt geringer als 70%, ganz besonders bevorzugt geringer als 60%, am meisten bevorzugt geringer als 50% über jeweils die gesamte Länge der transkribierten RNA oder mRNA.

"Zumindest teilweise doppelsträngiges Ribonukleinsäuremolekül" (infolge dsRNA) meint Ribonukleinsäuremolekül, die ganz oder teilweise doppelsträngig sind. Bevorzugt ist die Ribonukleinsäuresequenz überwiegend vollständig doppelsträngig. "Überwiegend vollständig doppelsträngig" meint, dass zumindest 50%, bevorzugt 70%, besonders bevorzugt 80%, ganz besonders bevorzugt 90% der in dem Molekül vorhandenen Basen in Paarung mit einer anderen Base der dsRNA vorliegen oder - entsprechend der Sequenz der dsRNA und den Basenpaarregeln sowie gegebenenfalls einer RNA-Sekundärstrukturvoraussage mittels eines geeigneten Computeralgorithmus - zumindest theoretisch vorliegen können.

"Im wesentlichen identisch" meint, dass eine "sense"-Ribonukleotidsequenz der dsRNA auch Insertionen, Deletionen sowie einzelne Punktmutationen im Vergleich zu der Sequenz des "sense"-RNA-Transkriptes eines endogenen Zielgens aufweisen kann. Mutationen umfassen Substitutionen, Additionen, Deletionen, Inversion oder Insertionen eines oder mehrerer Basen einer Nukleinsäuresequenz. Bevorzugt beträgt die Homologie zwischen einer "sense"-Ribonukleotidsequenz einer dsRNA und mindestens einem Teil des "sense"-RNA-Transkript eines endogenen Zielgens mindestens 60 %, bevorzugt mindestens 70 %, ganz besonders bevorzugt mindestens 90 %, am meisten bevorzugt 95%. Die Sequenzen können auch identisch mit der korrespondierenden Sequenz des Zielgens sein. Eine 100%ige Sequenzidentität zwischen der "sense"-Ribonukleotidsequenz der dsRNA und mindestens einem Teil des "sense"-Stranges der Transkriptes eines endogenen Gens ist bevorzugt, wenn gleich

nicht zwingend erforderlich, um eine effiziente Verminderung der Expression des endogenen Gens zu bewirken. Einzelne Mutationen werden toleriert. Das Verfahren ist demnach tolerant gegenüber Sequenzabweichungen, wie sie infolge genetischer Mutationen, Polymorphismen oder evolutionärer Divergenzen vorliegen können. So ist es beispielsweise auch möglich mit einer einzigen dsRNA, die ausgehend von einer bestimmten endogenen Gen generiert wurde, die Expression weiterer homologer endogener Gene des gleichen Organismus oder aber auch die Expression homologer endogener Gene in anderen verwandten Arten zu unterdrücken.

Unter Homologie wird das Maß der Übereinstimmung zwischen zwei Nukleotid-, Ribonukleotid- oder Proteinsequenzen verstanden, die bevorzugt durch Vergleich mit Hilfe des Programmalgorithmus GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA; Altschul et al. (1997) Nucleic Acids Res. 25:3389ff) unter Einstellung folgender Parameter berechnet wird:

20 Gap Weight: 50

Length Weight: 3

Average Match: 10

Average Mismatch: 0

Dem Fachmann ist bewusst, dass wenn die Homologie zwischen DNA (z.B. Genen) und RNA bestimmt wird, Thymin (T) in der DNA Sequenz als äquivalent zu Uracil (U) in der RNA Sequenz betrachtet wird.

"Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens" meint Fragmente einer RNA oder mRNA transkribiert von einem endogenen Zielgen. Dabei hat besagtes Teil bevorzugt eine Sequenzlänge von mindestens 10 Basen, bevorzugt mindestens 25 Basen, besonders bevorzugt mindestens 50 Basen, ganz besonders bevorzugt mindestens 100 Basen, am meisten bevorzugt mindestens 200 Basen oder mindestens 300 Basen. Umfasst ist auch die vollständige transkribierte RNA oder mRNA.

Alternativ, kann eine "im wesentlichen identische" dsRNA auch als Nukleinsäuresequenz definiert werden, die befähigt ist, mit einem Teil eines Transkriptes, bevorzugt der mRNA, eines endogenen Zielgenes zu hybridisieren (z.B. in 400 mM NaCl, 40 mM PIPES pH 6,4, 1 mM EDTA bei 50°C oder 70°C für 12 bis 16 h oder unter anderen Standardhybridisierungsbedingungen).

"Standardhybridisierungsbedingungen" ist breit zu verstehen und meint weniger stringente als auch - bevorzugt - stringente Hybridisierungsbedingungen. Solche Hybridisierungsbedingungen sind unter anderem bei Sambrook J, Fritsch EF, Maniatis T et al., in

Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57) oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. beschrieben.

5

Beispielhaft können die Bedingungen während des Waschschrilles ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit ungefähr 2X SSC bei 50°C) und - bevorzugt - solchen mit hoher Stringenz (mit ungefähr 0,2X SSC bei 50°C bevorzugt bei 65°C) (20X SSC: 0,3 M Natriumcitrat, 3 M NaCl, pH 7.0). Darüberhinaus kann die Temperatur während des Waschschrilles von niedrig stringenten Bedingungen bei Raumtemperatur, ungefähr 22°C, bis zu - bevorzugt - stärker stringenten Bedingungen bei ungefähr 65°C angehoben werden. Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50% Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt. Einige beispielhafte Bedingungen für Hybridisierung und Waschschrift sind in-  
folge gegeben:

(1) Hybridisierungsbedingungen zum Beispiel aus nachfolgenden Bedingungen ausgewählt sein:

- a) 4X SSC bei 65°C,
- b) 6X SSC bei 45°C,
- c) 6X SSC, 100 µg/ml denaturierter, fragmentierte Fischsperma-DNA bei 68°C,
- f) 50% Formamid, 4X SSC bei 42°C,
- h) 2X oder 4X SSC bei 50°C (schwach stringente Bedingung),
- i) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (schwach stringente Bedingung).

35

(2) Waschschrille können zum Beispiel aus nachfolgenden Bedingungen ausgewählt sein:

- a) 0,015 M NaCl/0,0015 M Natriumcitrat/0,1% SDS bei 50°C.
- b) 0,1X SSC bei 65°C.
- c) 0,1X SSC, 0,5% SDS bei 68°C.
- d) 0,1X SSC, 0,5% SDS, 50% Formamid bei 42°C.
- e) 0,2X SSC, 0,1% SDS bei 42°C.
- f) 2X SSC bei 65°C (schwach stringente Bedingung).

45

"Im wesentlichen komplementär" meint, dass die "antisense"-Ribonukleotidsequenzen der dsRNA auch Insertionen, Deletionen sowie

- einzelne Punktmutationen im Vergleich zu dem Komplement der "sense"-Ribonukleotidsequenzen aufweisen kann. Bevorzugt beträgt die Homologie mindestens 80 %, bevorzugt mindestens 90 %, ganz besonders bevorzugt mindestens 95 %, am meisten bevorzugt 100%
- 5 zwischen den "antisense"-Ribonukleotidsequenzen und dem Komplement der "sense"-Ribonukleotidsequenzen. Komplement meint dabei - in der dem Fachmann geläufigen Weise - den entsprechend den Basenpaarregeln abgeleiteten Gegenstrang.
- 10 Die doppelsträngige Struktur der dsRNA kann ausgehend von einem einzigen, ganz oder teilweise selbstkomplementären RNA-Strang (bei dem die oben erwähnten "sense"- und "antisense"-Ribonukleotidsequenzen der dsRNA alle kovalent miteinander verbunden sind) oder ausgehend von zwei RNA-Strängen (indem die oben erwähnten
- 15 "sense"- und "antisense"-Ribonukleotidsequenzen der dsRNA auf separate Stränge vorliegen), die zueinander ganz oder teilweise komplementär sind, gebildet werden. Bei zwei separaten Strängen können beispielsweise alle "sense"-Ribonukleotidsequenzen auf dem einen und alle "antisense"-Ribonukleotidsequenzen auf dem anderen
- 20 Strang vorliegen. Die Sequenzen können aber auch anders auf die beiden Stränge verteilt sein. Die Ausbildung der doppelsträngigen Struktur kann in vitro aber auch in vivo - beispielsweise in der eukaryotischen Zelle selber - erfolgen. Bevorzugt liegt die dsRNA in Form eines einzigen, selbstkomplementären RNA-Stranges vor.
- 25 Die einzelnen "sense"-Ribonukleotidsequenzen können mit den korrespondierenden, im wesentlichen komplementären "antisense"-Ribonukleotidsequenzen eine doppelsträngige RNA-Struktur mittels Basenpaarung ausbilden und bilden eine Untereinheit der dsRNA.
- 30 Im Falle eines selbstkomplementären Stranges ergeben sich verschiedene Möglichkeiten für die Primärstruktur der dsRNA. Nachfolgend aufgeführte sind beispielhaft, jedoch nicht einschränkend zu verstehen:
- 35 a) Es können zunächst die "sense"-Ribonukleotidsequenzen (S) der einzelnen Untereinheiten aneinander gefügt werden, worauf dann eine Aneinanderreihung der im wesentlichen komplementären "antisense"-Ribonukleotidsequenzen (AS) folgt. Die Anzahl der
- 40 Einheiten n ist größer oder gleich zwei. Es entsteht eine Struktur mit einer einzelnen Haarnadel. Die Primärstruktur der dsRNA kann dabei schematisch beispielsweise wie folgt aussehen:

45 5'-S(1)-S(2)-.....-S(n)-AS(n)-.....-AS(2)-AS(1)-3'

## 10

Die bevorzugte Sekundärstruktur ist in Fig. 2-A wiedergegeben.

- 5 b) Es können zunächst die "sense"-Ribonukleotidsequenz (S) und die im wesentlichen komplementäre "antisense"-Ribonukleotidsequenz (AS) der ersten Untereinheiten aneinander gefügt werden, wodrauf dann die Aneinanderreihung von "sense"- und "antisense"-Ribonukleotidsequenzen der weiteren Untereinheiten folgt. Die Anzahl der Einheiten n ist größer oder gleich
- 10 zwei. Es entsteht eine Struktur mit mehreren Haarnadeln. Die Primärstruktur der dsRNA kann dabei schematisch beispielsweise wie folgt aussehen:

5'-S(1)-AS(1)-S(2)-AS(2).....-S(n)-AS(n)-3'

15

Die bevorzugte Sekundärstruktur ist in Fig. 2-B wiedergegeben.

- Ist die dsRNA - bevorzugt - in der Lage eine Haarnadelstruktur auszubilden, so entspricht der Stamm der Haarnadel dem doppel-
- 20 strängige Anteil der dsRNA, der durch Basenpaarung zwischen auf dem gleich RNA-Moleküle lokalisierten "sense"- und "antisense"-Ribonukleotidsequenz gebildet wird. Dabei werden "sense"- und "antisense"-Ribonukleotidsequenzen bevorzugt durch einen "Linker" verbunden. Der "Linker" ist bevorzugt ein Intron, das aus der
- 25 dsRNA herausgespleißt werden kann. Selbstkomplementären dsRNA-Strukturen ausgehend von einem einzelnen RNA-Molekül sind bevorzugt, da sie lediglich die Expression eines Konstruktes erfordern und die komplementären RNA-Stränge stets in einem äquimolaren Verhältnis umfassen.

30

Bei der Verwendung eines Linkers (I) - bevorzugt eines Intron - seien nachfolgende schematische Primärstrukturen für die dsRNA beispielhaft genannt:

- 35 c) Dies ist eine bevorzugte Variante von a), bei der an der Stelle der Haarnadelschleife ein Linker (I) - bevorzugt ein Intron - insertiert wird:

5'-S(1)-S(2)-.....-S(n)-I-AS(n)-.....-AS(2)-AS(1)-3'

40

Die bevorzugte Sekundärstruktur ist in Fig. 2-C wiedergegeben.

- d) Dies ist eine bevorzugte Variante von b), bei der an der
- 45 Stelle der jeder Haarnadelschleife ein Linker (I) - bevorzugt ein Intron - insertiert wird:

5'-S(1)-I-AS(1)-S(2)-I-AS(2).....-S(n)-I-AS(n)-3'

Die bevorzugte Sekundärstruktur ist in Fig. 2-D wiedergegeben.

- Die dsRNA Moleküle sind jedoch auch ohne den Linker funktionell. Dabei ist jedoch zu berücksichtigen, dass die letzten ca. 10 Nukleotide der terminalen Untereinheit S(n) in diesem Fall nicht mehr korrekt paaren. In diesem Fall ist die Länge für diese Untereinheit um 10 Nukleotide zu ergänzen. Der Linker ist bevorzugt ein Intron, besonders bevorzugt ein Intron in sense-Orientierung. Bevorzugt handelt es sich um ein Intron eines pflanzlichen Gens. Beispielfhaft jedoch nicht einschränkend seien zu nennen: Das Intron 3 der Alkoholdehydrogenase 1 (Adh1) aus Mais (GenBank Acc.-No.: AF044293; GI: 2828164), das Intron 4 der beta-Conglycinin alpha Untereinheit aus Soja (GenBank Acc.-No.: AB051865); eines der Introns des rbcS-3A Gens für Ribulose-1.5-bisphosphatcarboxylase (RBC) kleine Untereinheit aus Erbse (GenBank Acc.-No.: X04333). Diese und weitere geeignete Introns sind dem Fachmann bekannt (McCullough AJ & Schuler MA (1997) Nuc Acids Res 25:1071-1077). Für die Anwendung in dem erfindungsgemäßen Verfahren wird das Intron bevorzugt in Kombination mit Spleißakzeptor- und Spleißdonorsequenzen eingesetzt, die ein späteres Heraus-spleißen aus der dsRNA ermöglichen. Diese Spleißsequenzen können die flankierenden Sequenzen des Intron selber sein, oder aber auch durch entsprechende Sequenzen der übrigen dsRNA bereitgestellt werden.
- Jede der einzelnen "sense"-Ribonukleotidsequenzen der dsRNA ist im wesentlichen identisch zu mindestens einem Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens. Dabei sind jedoch nicht alle "sense"-Ribonukleotidsequenzen zu dem "sense"-RNA-Transkript eines einzigen endogenen Zielgens identisch, sondern die jeweils maximale Identität von mindestens zwei der "sense"-Ribonukleotidsequenzen besteht zu den "sense"-RNA-Transkripten von unterschiedlichen endogenen Zielgenen. Dabei beträgt die Homologie zwischen den Transkripten der beiden endogenen Zielgene unter 90%, bevorzugt unter 80%, besonders bevorzugt unter 70%, ganz besonders bevorzugt unter 60%, am meisten bevorzugt unter 50%.

- Mindestens zwei der in der erfindungsgemäßen dsRNA umfassten einzelnen "sense"-Ribonukleotidsequenzen sind unterschiedlich. Unterschiedlich bedeutet zum einen, dass die Zielgene zu deren Transkripten sie die jeweils maximale Identität aufweisen, nicht identisch sind. Bevorzugt vermindert mindestens eine Untereinheit



## 12

- der dsRNA die Expression eines anderen Gens als mindestens eine andere Untereinheit. Unterschiedlich kann zum anderen auch heißen, dass die "sense"-Ribonukleotidsequenzen der Untereinheiten selber im wesentlichen nicht identisch sind und bevorzugt eine
- 5 Homologie zu einander unter 60%, besonders bevorzugt unter 50% ganz besonders bevorzugt unter 40% aufweisen. Die dsRNA kann in einer weiteren Ausführungsform mehrerer Kopien einer Untereinheit enthalten. Weiterhin kann die dsRNA auch mehrere verschiedene Untereinheiten enthalten, die aber gegen das gleiche endogene Ziel-
- 10 gens gerichtet sind und deren "sense"-Ribonukleotidsequenzen beispielsweise im wesentlichen identisch sind zu unterschiedlichen Teilen des "sense"-RNA-Transkriptes des besagten endogenen Zielgens.
- 15 Dabei kann jede der einzelnen "sense"-Ribonukleotidsequenzen auch zu dem Transkript mehrerer endogener Zielgene im wesentlichen identisch sein. Dies ist besonders dann der Fall, wenn die Zielgene über ähnliche Sequenzabschnitte verfügen, wie es beispielsweise bei Mitgliedern von Genfamilien (z.B. Speicherproteinen)
- 20 der Fall ist. Dies ist eine besonders vorteilhafte Anwendungsform, da - bei entsprechender Wahl der Ribonukleotidsequenz einer Untereinheit - besagte Untereinheit die Expression von mehr als einem Zielgen vermindern kann.
- 25 Vorzugsweise wird die Sequenz der dsRNA so gewählt, dass die angestrebte dsRNA Struktur nach Ausbildung der Duplex - im Vergleich zu anderen möglichen Faltungsvarianten der Primärstruktur der dsRNA - die jeweils geringste freie Energie hat. Dies kann beispielsweise durch Vermeidung von Sequenzduplikationen etc. ge-
- 30 währleistet werden. Die spezifische Sekundärstruktur kann beispielsweise mit geeigneten Computerprogrammen vorausgesagt und optimiert werden (z.B. FOLDRNA; Zuker and Stiegler (1981) Nucleic Acids Res 9(1):133-48).
- 35 Jede Untereinheit der dsRNA hat in einer bevorzugten Ausführungsform eine Länge von mindestens 20 Basenpaaren, bevorzugt mindestens 50 Basenpaaren, besonders bevorzugt mindestens 100 Basenpaare, ganz besonders bevorzugt mindestens 250 Basenpaare.
- 40 In einer weiterhin bevorzugten Ausführungsform hat jede Einheit eine Länge eine ganzzahligen Vielfachen von 21 oder 22 Basenpaaren, also beispielsweise 21, 22, 42, 43, 44, 63, 64, 65, 66, 84, 85, 86, 87, 88, 105, 106, 107, 108, 109, 110, 126, 127, 128, 129, 131, 132, 147, 148, 149, 150, 151, 152, 153, 154, 168, 169, 170,
- 45 171, 172, 173, 174, 175, 176, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219 oder 220 Basenpaare, bevorzugt 21, 22, 42, 44, 63, 66, 84, 88,

## 13

105, 110, 126, 132, 147, 154, 168, 176, 189, 198, 210 oder 220 Basenpaare, ganz besonders bevorzugt 21, 42, 63, 84, 105, 126, 147, 168, 189 oder 210 Basenpaare, am meisten bevorzugt 180 oder 210 Basenpaare.

5

Die "sense"- und/oder "antisense"-Ribonukleotidsequenzen der einzelnen Untereinheiten können direkt oder aber durch einen "Spacer" (SP; Abstandshalter) miteinander verbunden und/oder flankiert sein. Die einzelnen "Spacer" (SP) können dabei identisch  
 10 oder aber auch unterschiedlich sein. Der "Spacer" genügt dabei bevorzugt den gleichen Längenanforderungen wie sie oben für die Länge der Untereinheiten selber gegeben sind. Der "Spacer" kann eine doppelstränge Struktur ausbilden, kann aber auch - beispielsweise in Form einer Blase - in ungepaarter Formation bestehen, d.h. die Basen in Strang und Gegenstrang müssen nicht zwin-  
 15 genderweise komplementär sein. Bevorzugte Ausführungsformen sind zum Beispiel durch nachfolgende Primärstrukturen beschrieben:

e) Dies ist eine bevorzugte Variante von c):

20

5' SP-S(1)-SP-S(2)-SP-...-S(n)-AS(n)-SP-...-AS(2)-SP-AS(1)-SP-3'

Die bevorzugte Sekundärstruktur ist in Fig. 3-A wiedergegeben.

25

Der "Spacer" kann weitere Funktionselemente umfassen. Beispielfhaft jedoch nicht einschränkend sind zu nennen:

i) Sequenzen kodierend für eine von einem Ribozym als Substrat  
 30 erkannten Erkennungssequenz (RE). Beispielsweise kann die dsRNA nachfolgende lineare Struktur vor der Faltung einnehmen:

5'-S(1)-(RE)-S(2)-...-S(n)-AS(n)-...-AS(2)-(RE)-AS(1)-3'

35

Die bevorzugte Sekundärstruktur ist in Fig. 3-B wiedergegeben. Das entsprechende Ribozym (R) kann separat exprimiert werden kann aber auch auf der dsRNA selber kodiert sein. Dabei ist die Sequenz kodierend für ein Ribozym bevorzugt so  
 40 angeordnet, dass sie im gefalteten dsRNA Molekül einer Sequenz gegenüber liegt, die für dieses Ribozym als Substrat fungieren kann. Beispielsweise kann die dsRNA nachfolgende lineare Struktur vor der Faltung einnehmen:

45 5'-S(1)-(R)(RE)-S(2)-...-S(n)-AS(n)-...-AS(2)-(R)(RE)-AS(1)-3'

## 14

Die bevorzugte Sekundärstruktur ist in Fig. 3-C wiedergegeben. Durch die genannten Ausführungsformen werden nach Transkription die einzelnen Untereinheiten durch Wirkung des Ribozym voneinander getrennt. Diese Trennung ist vorteilhaft,  
5 jedoch nicht zwingend erforderlich. Entsprechend nutzbare Ribozyme und Erkennungssequenzen sind dem Fachmann bekannt.

Ribozyme meint katalytische RNA-Moleküle. Ribozyme können an jede beliebige Ziel-RNA angepasst werden und spalten das  
10 Phosphodiester-Gerüst an spezifischen Positionen, wodurch die Ziel-RNA funktionell deaktiviert wird (Tanner NK (1999) FEMS Microbiol Rev 23(3):257-275). Das Ribozym wird dadurch nicht selber modifiziert, sondern ist in der Lage, weitere Ziel-RNA-Moleküle analog zu spalten, wodurch es die Eigenschaften  
15 eines Enzyms erhält. Der Einbau von Ribozymsequenzen in "antisense"-RNAs verleiht eben diesen "antisense"-RNAs diese enzymähnliche, RNA-spaltende Eigenschaft und steigert so deren Effizienz bei der Inaktivierung der Ziel-RNA. Die Herstellung und Verwendung entsprechender Ribozym-"antisense"-RNA-Moleküle ist beispielsweise beschrieben bei Haseloff et al.  
20 (1988) Nature 334: 585-591. Auf diese Art können Ribozyme (z.B. "Hammerhead"-Ribozyme; Haselhoff und Gerlach (1988) Nature 334:585-591) verwendet werden, um die eines bestimmte RNA katalytisch zu spalten. Verfahren zur Expression von Ribozymen zur Verminderung bestimmter Proteine sind beschrieben  
25 in (EP 0 291 533, EP 0 321 201, EP 0 360 257). In pflanzlichen Zellen ist eine Ribozym-Expression ebenfalls beschrieben (Steinecke P et al. (1992) EMBO J 11(4):1525-1530; de Feyter R et al. (1996) Mol Gen Genet. 250(3):329-338). Geeignete  
30 Zielsequenzen und Ribozyme können zum Beispiel wie bei "Steinecke P, Ribozymes, Methods in Cell Biology 50, Galbraith et al. eds, Academic Press, Inc. (1995), S.449-460" beschrieben, durch Sekundärstrukturberechnungen von Ribozym- und Ziel-RNA sowie durch deren Interaktion bestimmt werden (Bayley CC et al. (1992) Plant Mol Biol. 18(2):353-361; Lloyd AM and Davis  
35 RW et al. (1994) Mol Gen Genet. 242(6):653-657). Beispielsweise können Derivate der Tetrahymena L-19 IVS RNA konstruiert werden, die komplementäre Bereiche zu der mRNA des zu den Spacersequenzen aufweisen (siehe auch US 4,987,071 und  
40 US 5,116,742). Alternativ können solche Ribozyme auch über einen Selektionsprozess aus einer Bibliothek diverser Ribozyme identifiziert werden (Bartel D und Szostak JW (1993) Science 261:1411-1418).

45 ii) Sequenzen kodierend für Erkennungssequenzen für RNAasen  
Der "Spacer" kann Erkennungssequenzen für RNAasen, bevorzugt sequenzspezifische RNAasen wie beispielsweise RNase III ent-

## 15

halten. RNase III schneidet am Motiv 5'-AGNN-3, wenn vier dieser Motive in einer Schleife vorhanden sind (Nagel R & Ares M (2000) RNA 6:1142-1156). Die RNase kann eine pflanzenneigene RNase sein, oder - wie beispielsweise für bakterielle RNase III Proteine - auch transgen exprimiert werden.

iii) Sequenzen kodierend für Intronspleißsignale (IS). Dabei sind die Spleißdonor und Spleißakzeptorsequenzen bevorzugt so lokalisiert, dass jeweils die Untereinheit als Intron herausgespleißt wird. Intronspleißsignale sind in Meritt et al. (1997) Plant Journal 12:937-943 oder in Egoavil et al. (1997) Plant Journal 12:971-980 beschrieben.

Die dsRNA bzw. ihre Vorläufermoleküle können auf verschiedene dem Fachmann geläufige Weise in einen Organismus oder eine Zelle eingebracht werden. "Einbringen" ist breit zu verstehen und umfasst im Rahmen der Erfindung alle Verfahren, die dazu geeignet eine dsRNA bzw. ihre Vorläufermoleküle, direkt oder indirekt, in einen Organismus oder eine Zelle, Kompartiment, Gewebe, Organ oder Samen desselben einzuführen oder dort zu generieren. Die Einbringung kann zu einer vorübergehenden (transienten) Präsenz einer dsRNA führen oder aber auch zu einer dauerhaften (stabilen). Umfasst sind Verfahren der direkten Transfektion oder Transformation der Zelle mit der als auch die Transformation oder Transfektion der Zelle mit Expressionskassetten, die befähigt sind, die der dsRNA zugrundeliegenden Ribonukleinsäuresequenzen in der Zelle zu exprimieren (infolge dsRNA-Expressionssystem). Die Expression der dsRNA kann transient oder - beispielsweise nach Integration in das Genom des Organismus - permanent erfolgen. Die Duplex-Bildung der dsRNA kann entweder außerhalb der Zelle oder innerhalb derselben initiiert werden.

Die dsRNA wird in einer Menge eingeführt, die zumindest eine Kopie pro Zelle ermöglicht. Höhere Mengen (z.B. mindestens 5, 10, 100, 500 oder 1000 Kopien pro Zelle) können ggf. eine effizientere Verminderung der Expression der Zielgene bewirken. Da dsRNA eine außerordentlich gute Mobilität innerhalb eines Organismus hat, ist es nicht zwingend erforderlich die dsRNA in jede Zelle des Organismus zu applizieren. Es ist ausreichend, die dsRNA in eine oder wenige Zellen einzubringen oder zu exprimieren, wobei die erfindungsgemäße Wirkung dann auch in anderen Zellen des gleichen Organismus erzielt werden kann.

Eine dsRNA - beispielsweise zur Verwendung in einer direkten Transformation oder Transfektion - kann in vivo oder in vitro, durch enzymatische, molekularbiologische oder chemisch-synthetische Verfahren synthetisiert werden. Dazu können eukaryoti-

sche, prokaryotische oder Bakteriophagen RNA Polymerasen (wie z.B. T3-, T7- oder SP6 RNA-Polymerase) verwendet werden. Entsprechende Verfahren zu in vitro Expression von RNA sind beschrieben (WO 97/32016; US 5,593,874; US 5,698,425, US 5,712,135, US 5,789,214, US 5,804,693). Eine chemisch oder enzymatisch in vitro synthetisierte dsRNA kann vor der Einführung in eine Zelle, Gewebe oder Organismus aus dem Reaktionsgemisch beispielsweise durch Extraktion, Präzipitation, Elektrophorese, Chromatographie oder Kombinationen dieser Verfahren ganz oder teilweise aufgereinigt werden. Die dsRNA kann direkt in die Zelle eingeführt werden (beispielsweise durch Partikelbeschuss oder Mikroinjektion) oder aber extrazellulär (z.B. in den interstitial Raum, das Gefäßsystem, das Verdauungssystem o.ä.) appliziert werden. Auch eine Applikation beispielsweise von dsRNA exprimierenden Organismen in Form von Nahrung ist denkbar. Es ist bekannt, dass dsRNA eine gute Zellgängigkeit und ausreichende Stabilität hat. Durch die hohe Wirksamkeit der dsRNA sind auch wenige Moleküle ausreichend, um eine gute Wirkung im Sinne der Erfindung zu erzielen.

Es können ferner Modifikationen sowohl des Zucker-Phosphat-Gerüsts als auch der Nukleoside in der dsRNA vorliegen. Beispielsweise können die Phosphodiesterbindungen der RNA dahingehend modifiziert sein, dass sie zumindest ein Stickstoff oder Schwefel-Heteroatom umfassen. Basen können dahingehend modifiziert werden, dass die Aktivität beispielsweise von Adenosindeaminase eingeschränkt wird. Die dsRNA kann enzymatisch oder ganz oder teilweise chemisch-synthetisch hergestellt werden.

Bevorzugt wird die dsRNA jedoch ausgehend von entsprechenden Expressionssystemen in der Zelle exprimiert. Ein weiterer Gegenstand der Erfindung betrifft besagte dsRNA-Expressionssysteme. Wird die dsRNA als ein einzelner, selbstkomplementärer RNA-Strang exprimiert, so umfasst das Expressionssystem eine Expressionskassette mit einer für den selbstkomplementären RNA-Strang kodierenden DNA Sequenz in funktioneller Verknüpfung mit einem Promotor, der geeignet ist, die Expression in der jeweiligen eukaryotischen Zelle zu gewährleisten. Optional kann die Expressionskassette weitere funktionelle Elemente wie beispielsweise Transkriptionsterminatoren und/oder Polyadenylierungssignale umfassen. Derartige Expressionskassetten sind ebenfalls Gegenstand der Erfindung.

Wird die dsRNA in Form von zwei separaten Strängen exprimiert, die zueinander ganz oder teilweise komplementär sind, so umfasst das Expressionssystem zwei Expressionskassetten, wobei jeder der beiden Stränge in funktioneller Verknüpfung mit einem Promotor steht, der geeignet ist, die Expression in der jeweiligen euka-

ryotischen Zelle zu gewährleisten. Optional können die Expressionskassetten weitere funktionelle Elemente wie beispielsweise Transkriptionsterminatoren und/oder Polyadenylierungssignale umfassen. Die Kombination der beiden Expressionskassetten zu dem  
5 erfindungsgemäßen Expressionssystem kann auf verschiedene dem Fachmann geläufige Art geschehen. Beispielfhaft seien zu nennen:

- a) Transformation der Zelle oder Pflanze mit einem Vektor, der Expressionskassetten für beide RNA-Stränge umfasst,  
10
- b) Kotransformation der Zelle oder Pflanze mit zwei Vektoren, wobei jeweils ein Vektor für jeweils einen der beiden Stränge der dsRNA kodiert.
- 15 c) Kreuzung von zwei Pflanzen, die mit jeweils einem Vektor transformiert wurden, wobei jeweils ein Vektor für jeweils einen der beiden Stränge der dsRNA kodiert.

Es ist auch möglich, das eine Expressionskassette einzusetzen,  
20 bei der die für die dsRNA kodierende DNA-Sequenz zwischen zwei Promotoren mit entgegengerichteter Transkriptionsrichtung lokalisiert ist und so von beiden Seiten transkribiert wird.

Expressionskassette meint chimäre DNA-Moleküle in denen eine für  
25 das dsRNA-Molekül (bzw. für einen der Stränge desselben) kodierende Nukleinsäuresequenz mit mindestens einem genetischen Kontrollelement (beispielsweise einem Promotor, Enhancer, Silencer, Splice-Donor oder -Akzeptor, Polyadenylierungssignal) derart verknüpft ist, das die Transkription des dsRNA-Moleküls (bzw. eines  
30 der Stränge desselben) in der eukaryotischen Zelle oder Organismus gewährleistet ist. Entsprechend vorteilhafte Konstruktionen sind weiter unten beschrieben. Eine Polyadenylierung ist möglich, jedoch nicht erforderlich, ebenso müssen keine Elemente zur Initiierung einer Translation vorhanden sein.

35 Soll das Expressionskonstrukt in eine Pflanze eingeführt und die dsRNA in planta erzeugt werden, so sind pflanzenspezifische genetische Kontrollelemente (beispielsweise pflanzenspezifische Promotoren) bevorzugt. Die dsRNA kann jedoch auch in anderen Or-  
40 ganismen oder in vitro erzeugt und dann in die Pflanze eingebracht werden.

Unter einer funktionellen Verknüpfung versteht man zum Beispiel die sequentielle Anordnung eines Promotors mit der zu transkri-  
45 bierenden Nukleinsäuresequenz und ggf. weiterer regulativer Elemente wie zum Beispiel einem Terminator und/oder Polyadenylierungssignalen derart, dass jedes der regulativen Elemente seine

## 18

Funktion bei der Transkription der Nukleinsäuresequenz, je nach Anordnung der Nukleinsäuresequenzen erfüllen kann. Dazu ist nicht unbedingt eine direkte Verknüpfung im chemischen Sinne erforderlich. Genetische Kontrollsequenzen, wie zum Beispiel Enhancer-Sequenzen, können ihre Funktion auch von weiter entfernten Positionen oder gar von anderen DNA-Molekülen aus auf die Zielsequenz ausüben. Bevorzugt sind Anordnungen, in denen die zu transkribierende Nukleinsäuresequenz hinter der als Promoter fungierenden Sequenz positioniert wird, so dass beide Sequenzen kovalent miteinander verbunden sind. Bevorzugt ist dabei der Abstand zwischen der Promotorsequenz und der transgen zu exprimierende Nukleinsäuresequenz geringer als 200 Basenpaare, besonders bevorzugt kleiner als 100 Basenpaare, ganz besonders bevorzugt kleiner als 50 Basenpaare. In einer bevorzugten Ausführungsform wird die zu transkribierende Nukleinsäuresequenz so hinter dem Promotor lokalisiert, dass der Transkriptionsstart identisch ist mit dem gewünschten Beginn der dsRNA.

Die Herstellung einer funktionellen Verknüpfung als auch die Herstellung einer Expressionskassette kann mittels gängiger Rekombinations- und Klonierungstechniken realisiert werden, wie sie beispielsweise in Maniatis T, Fritsch EF und Sambrook J (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), in Silhavy TJ, Berman ML und Enquist LW (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), in Ausubel FM et al. (1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience und bei Gelvin et al. (1990) In: Plant Molecular Biology Manual beschrieben sind.

Unter einer Expressionskassette sind aber auch solche Konstruktionen zu verstehen, bei denen zum Beispiel eine Nukleinsäuresequenz kodierend für eines dsRNA derart hinter einen endogenen Promotor platziert werden, dass der gleiche Effekt auftritt. Beide Ansätze führen zu Expressionskassetten im Sinne der Erfindung.

Prinzipiell können alle natürlichen Promotoren mit ihren Regulationssequenzen wie die oben genannten für das erfindungsgemäße Verfahren verwendet werden, solange sie die Expression in dem Zielorganismus gewährleisten. Darüberhinaus können auch synthetische Promotoren vorteilhaft verwendet werden.

Es können weitere Promotoren funktionell mit der zu exprimierenden Nukleinsäuresequenz verknüpft sein, die eine Expression in weiteren Eukaryoten oder in Prokaryoten, wie zum Beispiel *E.coli* Bakterien ermöglichen.

5

Die in den erfindungsgemässen Expressionskassetten oder Vektoren enthaltenen Nukleinsäuresequenzen können mit weiteren genetischen Kontrollsequenzen neben einem Promoter funktionell verknüpft sein. Der Begriff der genetischen Kontrollsequenzen ist breit zu  
10 verstehen und meint all solche Sequenzen, die einen Einfluss auf das Zustandekommen oder die Funktion der erfindungsgemässen Expressionskassette haben. Genetische Kontrollsequenzen modifizieren zum Beispiel die Transkription in prokaryotischen oder eukaryotischen Organismen. Vorzugsweise umfassen die erfindungsgemässen  
15 Expressionskassetten 5'-stromaufwärts von der jeweiligen transgen zu exprimierenden Nukleinsäuresequenz einen pflanzenspezifischen Promoter und 3'-stromabwärts eine Terminatorsequenz als zusätzliche genetische Kontrollsequenz, sowie gegebenenfalls weitere übliche regulative Elemente, und zwar jeweils funktionell  
20 verknüpft mit der transgen zu exprimierenden Nukleinsäuresequenz.

Genetische Kontrollsequenzen umfassen ferner auch die 5'-untranslatierte Regionen, Introns oder nichtkodierende 3'-Region von Genen. Es ist gezeigt worden, dass diese eine signifikante Funktion  
25 bei der Regulation der Genexpression spielen können. Kontrollsequenzen umfassen ferner Polyadenylierungssignale sowie Terminatorsequenzen.

Die Expressionskassette kann vorteilhafterweise eine oder mehrere  
30 sogenannte "enhancer Sequenzen" funktionell verknüpft mit dem Promoter enthalten, die eine erhöhte transgene Expression der Nukleinsäuresequenz ermöglichen. Auch am 3'-Ende der transgen zu exprimierenden Nukleinsäuresequenzen können zusätzliche vorteilhafte Sequenzen inseriert werden, wie weitere regulatorische Elemente oder Terminatoren. Die transgen zu exprimierenden Nukleinsäuresequenzen können in einer oder mehreren Kopien im Genkonstrukt enthalten sein.  
35

Als Kontrollsequenzen sind weiterhin solche zu verstehen, die  
40 eine homologe Rekombination bzw. Insertion in das Genom eines Wirtsorganismus ermöglichen oder die Entfernung aus dem Genom erlauben. Methoden wie die cre/lox-Technologie erlauben eine gewebespezifische, unter Umständen induzierbare Entfernung der Expressionskassette aus dem Genom des Wirtsorganismus (Sauer B  
45 (1998) Methods. 14(4):381-92). Hier werden bestimmte flankierende



Sequenzen dem Zielgen angefügt (lox-Sequenzen), die später eine Entfernung mittels der cre-Rekombinase ermöglichen.

Bevorzugt kann die Expressionskassette, bestehend aus einer Ver-  
5 knüpfung von Promoter und zu transkribierender Nukleinsäurese-  
quenz, integriert in einem Vektor vorliegen und durch zum Bei-  
spiel Transformation - nach einem der unten beschriebenen Verfah-  
ren - in die eukaryotische Zelle oder Organismus eingebracht wer-  
den. Die nachfolgende Expression kann transient sein oder aber  
10 auch - bevorzugt - stabil nach Insertion (beispielsweise unter  
Verwendung von Selektionsmarkern) der Expressionskassetten in das  
Genom erfolgen. Bevorzugt wird das dsRNA-Expressionssystem stabil  
in das Genom - beispielsweise die chromosomale DNA oder die DNA  
der Organellen (z.B. der Plastiden, Mitochondrien etc.) - einer  
15 Zelle integriert.

Die Einführung einer erfindungsgemässen transgenen Expressions-  
kassette in einen Organismus oder Zellen, Gewebe, Organe, Teile  
bzw. Samen desselben (bevorzugt in Pflanzen bzw. pflanzliche Zel-  
20 len, Gewebe, Organe, Teile oder Samen) kann vorteilhaft unter  
Verwendung von Vektoren realisiert werden, in denen die transge-  
nen Expressionskassetten enthalten sind. Vektoren können bei-  
spielhaft Plasmide, Cosmide, Phagen, Viren oder auch Agrobakte-  
rien sein. Die transgenen Expressionskassetten können in den Vek-  
25 tor (bevorzugt ein Plasmidvektor) über eine geeignete Restrikti-  
onsschnittstelle insertiert werden. Der entstandene Vektor wird  
zunächst in E.coli eingeführt. Korrekt transformierte E.coli wer-  
den selektioniert, gezüchtet und der rekombinante Vektor mit dem  
Fachmann geläufigen Methoden gewonnen. Restriktionsanalyse und  
30 Sequenzierung können dazu dienen, den Klonierungsschritt zu über-  
prüfen. Bevorzugt sind solche Vektoren, die eine stabile Integra-  
tion der Expressionskassette in das Wirtsgenom ermöglichen.

Die Herstellung eines transformierten Organismus (bzw. einer  
35 transformierten Zelle oder Gewebes) erfordert, dass die entspre-  
chende DNA (z.B. der Expressionsvektor) oder RNA in die entspre-  
chende Wirtszelle eingebracht wird. Für diesen Vorgang, der als  
Transformation (oder Transduktion bzw. Transfektion) bezeichnet  
wird, steht eine Vielzahl von Methoden zur Verfügung (Keown et  
40 al. (1990) Methods in Enzymology 185:527-537). So kann die DNA  
oder RNA beispielhaft direkt durch Mikroinjektion oder durch Bom-  
bardierung mit DNA-beschichteten Mikropartikeln eingeführt wer-  
den. Auch kann die Zelle chemisch, zum Beispiel mit Polyethylen-  
glycol, permeabilisiert werden, so dass die DNA durch Diffusion  
45 in die Zelle gelangen kann. Die DNA kann auch durch Protopla-  
stenfusion mit anderen DNA-enthaltenden Einheiten wie Minicells,  
Zellen, Lysosomen oder Liposomen erfolgen. Elektroporation ist

## 21

eine weitere geeignete Methode zur Einführung von DNA, bei der die Zellen reversibel durch einen elektrischen Impuls permeabilisiert werden. Entsprechende Verfahren sind beschrieben (beispielsweise bei Bilanz et al. (1991) *Gene* 100:247-250; Scheid et al. (1991) *Mol Gen Genet* 228:104-112; Guerche et al. (1987) *Plant Science* 52:111-116; Neuhaus et al. (1987) *Theor Appl Genet* 75:30-36; Klein et al. (1987) *Nature* 327:70-73; Howell et al. (1980) *Science* 208:1265; Horsch et al. (1985) *Science* 227:1229-1231; DeBlock et al. (1989) *Plant Physiology* 91:694-701; 10 *Methods for Plant Molecular Biology* (Weissbach and Weissbach, eds.) Academic Press Inc. (1988); and *Methods in Plant Molecular Biology* (Schuler and Zielinski, eds.) Academic Press Inc. (1989)).

15 Ein weiterer Gegenstand der Erfindung betrifft Zellen, die eines der erfindungsgemäßen dsRNA Moleküle, Expressionssysteme, Expressionsskassetten oder Expressionsvektoren enthalten. Die Zelle kann von einem Organismus abgeleitet oder in diesem enthalten sein, meint aber auch einzellige Organismen wie Mikroorganismen. Die 20 Zelle kann prokaryotisch oder eukaryotischer Natur sein. Wobei das erfindungsgemäße Verfahren auf eukaryotische Organismen angewendet wird. Dennoch können prokaryotische Organismen die erfindungsgemäßen Expressionssysteme beispielsweise zum Zwecke der dsRNA-Produktion enthalten. Auch können prokaryotische Organismen, beispielsweise Agrobakterien, vorteilhaft als Vehikel für 25 die Transformation beispielsweise pflanzlicher Organismen eingesetzt werden.

Bevorzugte Prokaryoten sind vor allem Bakterien wie Bakterien der 30 Gattung *Escherichia*, *Corynebacterium*, *Bacillus*, *Clostridium*, *Proionibacterium*, *Butyrivibrio*, *Eubacterium*, *Lactobacillus*, *Erwinia*, *Agrobacterium*, *Flavobacterium*, *Alcaligenes*, *Phaeodactylum*, *Colpidium*, *Mortierella*, *Entomophthora*, *Mucor*, *Cryptocodium* oder Cyanobakterien zum Beispiel der Gattung *Synechocystis*. Bevorzugt sind vor allem Mikroorganismen, welche zur Infektion von 35 Pflanzen und damit zur Übertragung der erfindungsgemäßen Konstrukte befähigt sind. Bevorzugte Mikroorganismen sind solche aus der Gattung *Agrobacterium* und insbesondere der Art *Agrobacterium tumefaciens*.

40

Eukaryotische Zellen und Organismen umfasst pflanzliche und tierische, nicht-menschliche Organismen und/oder Zellen sowie eukaryotische Mikroorganismen wie beispielsweise Hefen, Algen oder Pilze. Ein entsprechender transgener Organismus kann beispielsweise 45 durch Einführung der entsprechenden Expressionssysteme in

eine Zygote, Stammzelle, Protoplast oder eine andere geeignete von dem Organismus abgeleitete Zelle hergestellt werden.

"Tierische Organismus" meint nicht-menschliche Vertebraten oder  
5 Invertebraten. Bevorzugte Vertebraten umfassen beispielsweise Fischearten, nicht-menschliche Säugetiere wie Rind, Pferd, Schaf, Ziege, Maus, Ratte oder Schwein, sowie Vögel wie Huhn oder Gans. Bevorzugte tierische Zellen umfassen CHO, COS, HEK293 Zellen. Invertebraten umfassen Nematoden oder andere Würmer sowie Insekten.  
10 Invertebraten umfassen Insektenzellen wie Drosophila S2 und Spodoptera Sf9 oder Sf21 Zellen.

Bevorzugt sind ferner Nematoden, die in der Lage sind Tiere oder Menschen zu befallen, wie solche der Gattungen Ancylostoma, Ascaridia, Ascaris, Bunostomum, Caenorhabditis, Capillaria, Chabertia, Cooperia, Dictyocaulus, Haemonchus, Heterakis, Nematodirus, Oesophagostomum, Ostertagia, Oxyuris, Parascaris, Strongylus, Toxascaris, Trichuris, Trichostrongylus, Trichonema, Toxocara oder Uncinaria. Ferner bevorzugt sind solche, die in der Lage sind  
20 pflanzliche Organismen zu befallen wie beispielsweise Bursaphelenchus, Criconemella, Diitylenchus, Ditylenchus, Globodera, Helicotylenchus, Heterodera, Longidorus, Meloidogyne, Nacobbus, Pratylenchus, Pratylenchus, Radopholus, Rotelynchus, Tylenchus oder Xiphinema. Bevorzugte Insekten umfassen solche der Gattungen Coleoptera, Diptera, Lepidoptera und Homoptera.  
25

Bevorzugte Pilze sind Aspergillus, Trichoderma, Ashbya, Neurospora, Fusarium, Beauveria oder weitere in Indian Chem Engr. Section B. Vol 37, No 1,2 (1995) auf Seite 15, Tabelle 6 beschriebene Pilze. Besonders bevorzugt ist der filamentöse Hemiascomycet Ashbya gossypii.  
30

Bevorzugte Hefen sind Candida, Saccharomyces, Hansenula oder Pichia, besonders bevorzugt sind Saccharomyces cerevisiae oder Pichia pastoris (ATCC Accession No. 201178).  
35

Als transgene Organismen bevorzugt sind vor allem pflanzliche Organismen. "Pflanzlicher Organismus" umfasst jeden Organismus, der zur Photosynthese befähigt ist, sowie die von diesem abgeleitete  
40 Zellen, Gewebe, Teile oder Vermehrungsgut (wie Samen oder Früchte). Eingeschlossen sind im Rahmen der Erfindung alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches. Einjährige, mehrjährige, monocotyledone und dicotyledone Pflanzen sowie Gymnospermen sind bevorzugt. Eingeschlossen sind  
45 reife Pflanze, Saatgut, Sprosse und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut (zum Beispiel Knollen, Samen oder Früchte) und Kulturen, zum Beispiel Zell- oder Kalluskulturen.

## 23

Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium.

- 5 "Pflanze" im Rahmen der Erfindung meint alle Gattungen und Arten höherer und niedrigerer Pflanzen des Pflanzenreiches. Eingeschlossen unter dem Begriff sind die reifen Pflanzen, Saatgut, Sprossen und Keimlinge, sowie davon abgeleitete Teile, Vermehrungsgut, Pflanzenorgane, Gewebe, Protoplasten, Kallus und andere
- 10 Kulturen, zum Beispiel Zellkulturen, sowie alle anderen Arten von Gruppierungen von Pflanzenzellen zu funktionellen oder strukturellen Einheiten. Reife Pflanzen meint Pflanzen zu jedem beliebigen Entwicklungsstadium jenseits des Keimlings. Keimling meint eine junge, unreife Pflanze in einem frühen Entwicklungsstadium.
- 15 "Pflanze" umfasst alle einjährigen und mehrjährige, monokotyledonen und dikotyledonen Pflanzen und schließt beispielhaft jedoch nicht einschränkend solche der Gattungen Cucurbita, Rosa, Vitis, Juglans, Fragaria, Lotus, Medicago, Onobrychis, Trifolium, Trigonella, Vigna, Citrus, Linum, Geranium, Manihot, Daucus, Arabidopsis, Brassica, Raphanus, Sinapis, Atropa, Capsicum, Datura, Hyoscyamus, Lycopersicon, Nicotiana, Solarium, Petunia, Digitalis, Majorana, Cichorium, Helianthus, Lactuca, Bromus, Asparagus, Antirrhinum, Heterocallis, Nemesis, Pelargonium, Panieum, Pennisetum, Ranunculus, Senecio, Salpiglossis, Cucumis, Browaalia, Glycine, Pisum, Phaseolus, Lolium, Oryza, Zea, Avena, Hordeum, Secale, Triticum, Sorghum, Picea und Populus ein.

Bevorzugt sind Pflanzen nachfolgender Pflanzenfamilien: Amarantaceae, Asteraceae, Brassicaceae, Carophyllaceae, Chenopodiaceae,

30 Compositae, Cruciferae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanacea, Sterculiaceae, Tetragnoniacea, Theaceae, Umbelliferae.

- 35 Bevorzugte monokotyle Pflanzen sind insbesondere ausgewählt aus den monokotylen Kulturpflanzen, wie zum Beispiel der Familie der Gramineae wie Alfalfa, Reis, Mais, Weizen oder andere Getreidearten wie Gerste, Hirse, Roggen, Triticale oder Hafer sowie dem Zuckerrohr sowie alle Arten von Gräsern.

40

Die Erfindung wird ganz besonders bevorzugt aus dikotyledone pflanzliche Organismen angewendet. Bevorzugte dikotyle Pflanzen sind insbesondere ausgewählt aus den dikotylen Kulturpflanzen, wie zum Beispiel

45

## 24

- Asteraceae wie Sonnenblume, Tagetes oder Calendula und andere mehr,
  - Compositae, besonders die Gattung Lactuca, ganz besonders die  
5 Art sativa (Salat) und andere mehr,
  - Cruciferae, besonders die Gattung Brassica, ganz besonders die Arten napus (Raps), campestris (Rübe), oleracea cv Tastie (Kohl), oleracea cv Snowball Y (Blumenkohl) und oleracea cv Emperor (Broccoli) und weitere Kohlarten; und der Gattung Arabidopsis, ganz besonders die Art thaliana sowie Kresse oder Canola und andere mehr,  
10
  - Cucurbitaceae wie Melone, Kürbis oder Zucchini und andere mehr,  
15
  - Leguminosae besonders die Gattung Glycine, ganz besonders die Art max (Sojabohne) Soja sowie Alfalfa, Erbse, Bohnengewächsen oder Erdnuss und andere mehr
  - 20 - Rubiaceae, bevorzugt der Unterklasse Lamiidae wie beispielsweise Coffea arabica oder Coffea liberica (Kaffestrauch) und andere mehr,
  - Solanaceae besonders die Gattung Lycopersicon, ganz besonders  
25 die Art esculentum (Tomate) und die Gattung Solanum, ganz besonders die Art tuberosum (Kartoffel) und melongena (Aubergine) sowie Tabak oder Paprika und andere mehr,
  - Sterculiaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Theobroma cacao (Kakaostrauch) und andere mehr,  
30
  - Theaceae, bevorzugt der Unterklasse Dilleniidae wie beispielsweise Camellia sinensis oder Thea sinensis (Teestrauch) und andere mehr,  
35
  - Umbelliferae, besonders die Gattung Daucus (ganz besonders die Art carota (Karotte)) und Apium (ganz besonders die Art graveolens dulce (Selderie)) und andere mehr; und die Gattung Capsicum, ganz besonders die Art annum (Pfeffer) und andere mehr,  
40
- sowie Lein, Soya, Baumwolle, Hanf, Flachs, Gurke, Spinat, Möhre, Zuckerrübe und den verschiedenen Baum-, Nuss- und Weinarten, insbesondere Banane und Kiwi.
- 45 Umfasst sind ferner Schmuckpflanzen, Nutz- oder Zierbäume, Blumen, Schnittblumen, Sträucher oder Rasen. Beispielhaft aber nicht einschränkend seien zu nennen Angiospermen, Bryophyten wie zum

## 25

Beispiel Hepaticae (Leberblümchen) und Musci (Moose); Pteridophyten wie Farne, Schachtelhalm und Lycopoden; Gymnospermen wie Koniferen, Cycaden, Ginkgo und Gnetalen, die Familien der Rosaceae wie Rose, Ericaceae wie Rhododendrons und Azaleen, Euphorbiaceae wie Weihnachtssterne und Kroton, Caryophyllaceae wie Nelken, Solanaceae wie Petunien, Gesneriaceae wie das Usambaraveilchen, Balsaminaceae wie das Springkraut, Orchidaceae wie Orchideen, Iridaceae wie Gladiolen, Iris, Freesie und Krokus, Compositae wie Ringelblume, Geraniaceae wie Geranien, Liliaceae wie der Drachenbaum, Moraceae wie Ficus, Araceae wie Philodendron und andere mehr.

Pflanzliche Organismen im Sinne der Erfindung sind weiterhin weitere photosynthetisch aktive befähigte Organismen, wie zum Beispiel Algen, Cyanobakterien sowie Moose. Bevorzugte Algen sind Grünalgen, wie beispielsweise Algen der Gattung Haematococcus, Phaedactylum tricornatum, Volvox oder Dunaliella. Insbesondere bevorzugt ist Synechocystis.

20 Am meisten bevorzugt sind

- a) Pflanzen, die zur Ölproduktion geeignet sind, wie beispielsweise Raps, Sonnenblume, Sesam, Färberdistel (*Carthamus tinctorius*), Ölbaum, Soja, Mais, Erdnuß, Rizinus, Ölpalme, Weizen, Kakaostrauch oder verschiedene Nussarten wie beispielsweise Walnuss, Kokosnuß oder Mandel. Unter diesen wieder besonders bevorzugt sind dikotyledonen Pflanzen, insbesondere Raps, Soja und Sonnenblume.
- 30 b) Pflanzen, die der Stärkeproduktion dienen, wie beispielsweise Mais, Weizen oder Kartoffel.
- c) Pflanzen, die als Nahrungs- und/oder Futtermittel und/oder Nutzpflanze genutzt werden und bei denen eine Resistenz gg. Pathogene vorteilhaft wäre, wie beispielsweise Gerste, Roggen, Reis, Kartoffel, Baumwolle, Flachs, Lein.
- 35 d) Pflanzen, die zur Produktion von Feinchemikalien wie beispielsweise Vitaminen und/oder Carotinoiden dienen können, wie beispielsweise Raps.
- 40

Die in den Verfahren verwendeten Organismen werden je nach Wirtorganismus in dem Fachmann bekannter Weise angezogen bzw. gezüchtet. Mikroorganismen werden in der Regel in einem flüssigen Medium, das eine Kohlenstoffquelle meist in Form von Zuckern, eine Stickstoffquelle meist in Form von organischen Stickstoffquellen wie Hefeextrakt oder Salzen wie Ammoniumsulfat, Spurenelemente

wie Eisen-, Mangan-, Magnesiumsalze und gegebenenfalls Vitamine enthält, bei Temperaturen zwischen 0°C und 100°C, bevorzugt zwischen 10°C bis 60°C unter Sauerstoffbegasung angezogen. Dabei kann der pH der Nährflüssigkeit auf einen festen Wert gehalten werden, das heißt während der Anzucht reguliert werden oder nicht. Die Anzucht kann batch weise, semi batch weise oder kontinuierlich erfolgen. Nährstoffe können zu beginn der Fermentation vorgelegt oder semikontinuierlich oder kontinuierlich nach gefüttert werden.

10

Nachfolgende Anwendung des erfindungsgemäßen Verfahrens seien beispielhaft jedoch nicht einschränkend zu nennen:

#### I. Pflanzenbiotechnologie

15

Bevorzugt wird das erfindungsgemäße Verfahren im Rahmen der Pflanzenbiotechnologie zur Erzeugung von Pflanzen mit vorteilhaften Eigenschaften eingesetzt. So kann Eignung der Pflanzen oder deren Samen als Nahrungs- oder Futtermittel verbessert werden, beispielsweise über eine Veränderung der Zusammensetzungen und/oder des Gehalt an Metaboliten, insbesondere Proteinen, Ölen, Vitaminen und/oder Stärke. Auch können Wachstumsrate, Ertrag oder die Resistenz gegen biotische oder abiotische Stressfaktoren erhöht werden. Nachfolgende Anwendungen im Bereich der Pflanzenbiotechnologie sind insbesondere vorteilhaft. Die angegebenen möglichen Zielgene sind beispielhaft jedoch nicht einschränkend zu verstehen:

1. Verbesserter Schutz gegen abiotische Stressfaktoren (Hitze, Kälte, Trockenheit, erhöhte Feuchtigkeit, Umweltgifte, UV-Strahlung). Bevorzugt werden Gene in ihrer Expression vermindert, die an Stressreaktionen beteiligt sind.

30

2. Modifikation der Zusammensetzung und/oder des Gehaltes an Fettsäuren, Lipiden oder Ölen

35

Eine Veränderung des Fettsäuregehaltes oder der Fettsäurezusammensetzung, vorzugsweise in einer Ölfrucht wie Raps oder Sonnenblume, kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen der Fettsäurebiosynthese vorzugsweise ausgewählt aus der Gruppe bestehend aus Genen kodierend für Acetyltransacylasen, Acyltransportproteinen ("acyl carrier protein"), Desaturasen wie Stearyl-desaturasen oder mikrosomale  $\Delta 12$ -Desaturasen insbesondere Fad2-1 Gene, Malonyltransacylase,  $\beta$ -Ketoacyl-ACP-synthetasen, 3-Keto-ACP-reduktasen, Enoyl-ACP-hydrasen, Thioesterasen wie Acyl-ACP-thioesterases, Enoyl-ACP-reduktasen. Verschiedene weitere vor-

40

45

teilhafte Ansätze zur Modifizierung der Lipidzusammensetzung sind beschrieben (Shure M et al. (1983) Cell 35:225-233; Preiss et al. (1987) Tailoring Genes for Crop Improvement (Bruening et al., eds.), Plenum Press, S.133-152; Gupta et al. (1988) Plant Mol Biol. 10:215-224; Olive et al. (1989) Plant Mol Biol 12:525-538; Bhattacharyya et al. (1990) Cell 60:155-122; Dunwell JM (2000) J Exp Botany 51Spec No:487-96; Brar DS et al. (1996) Biotech Genet Eng Rev 13:167-79; Kishore GM und Somerville CR (1993) Curr Opin Biotech 4(2):152-8). Bevorzugt sind insbesondere Fad2 Gene (z.B. beschrieben durch Genbank Acc.-Nr.: AF124360 (*Brassica carinata*), AF042841 (*Brassica rapa*), L26296 (*Arabidopsis thaliana*), A65102 (*Corylus avellana*)). Weitere vorteilhafte Gene und Verfahren zur Modifikation des Lipidgehaltes sind beispielsweise beschrieben in US 5,530,192 und WO 94/18337. Ein erhöhter Lipidgehalt kann auch erreicht werden durch Verminderung des Stärkegehaltes beispielsweise infolge verminderter Expression von von Enzymen des Kohlenhydratstoffwechsels (z.B. ADP-Glucosepyrophosphorylasen).

### 3. Modifikation der Kohlenhydratzusammensetzung

Eine Modifikation der Kohlenhydratzusammensetzung kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen des Kohlenhydratstoffwechsels oder der Kohlenhydratbiosynthese, beispielsweise der Biosynthese von Amylose, Pektinen, Cellulose oder Zellwandkohlenhydraten. Dadurch kann eine Vielzahl zellulärer Prozesse (Reifung, Halfestigkeit, Stärkezusammensetzung oder -gehalt etc.) in vorteilhafter Weise beeinflusst werden. Als Zielgene seien beispielsweise jedoch nicht einschränkend zu nennen Phosphorylasen, Stärkesynthetasen, Q-Enzyme, Sucrose-6-phosphatsynthetasen, Sucrose-6-phosphatphosphatasen, ADP-Glucosepyrophosphorylasen, Branching-Enzyme, Debranching-Enzyme sowie diverse Amylasen. Die entsprechenden Gene sind beschrieben (Dunwell JM (2000) J Exp Botany 51Spec No:487-96; Brar DS et al. (1996) Biotech Genet Eng Rev 13:167-79; Kishore GM und Somerville CR (1993) Curr Opin Biotech 4(2):152-8). Vorteilhafte Gene zur beeinflussung des Kohlenhydratstoffwechsels - insbesondere der Stärkebiosynthese - sind beschrieben in WO 92/11375, WO 92/11376, US 5, 365,016 und WO 95/07355.

### 4. Veränderung der Farbe oder Pigmentierung

Veränderung der Farbe oder Pigmentierung vorzugsweise von Zierpflanzen kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen der Flavonoid-Biosynt-



hese wie beispielsweise Chalconsynthasen, Chalconisomerasen, Phenylalaninammonialyasen, Dehydrokaempferol(flavone)hydroxylasen wie Flavanon-3-hydroxylasen oder Flavanon-2-hydroxylasen, Dihydroflavonolreduktasen, Dihydroflavanol-2-hydroxylasen, Flavonoid-3'-hydroxylasen, Flavonoid-5'-hydroxylasen, Flavonoidglycosyltransferasen (z.B. Glucosyltransferasen wie UDPG:Flavonoid- 3-O-glucosyltransferasen, UDPG:Flavanol-7-O-glucosyltransferasen oder Rhamnosyltransferasen), Flavonoidmethyltransferasen (wie z.B. SAM:Anthocyanidin-3-(p-coumaroyl)-rutinosid-5-glucosid-3',5'-O-methyltransferasen) und Flavonoidacyltransferasen (Hahlbrock (1981) Biochemistry of Plants, Vol.7, Conn (Ed.); Weiring and de Vlamming (1984) "Petunia", KC Sink (Ed.), Springer-Verlag, New York). Geeignet sind insbesondere die in EP-A1 522 880 beschriebenen Sequenzen.

#### 5. Verminderung des Gehaltes von Speicherproteinen

Die Verminderung der Genexpression von Genen kodierend für Speicherproteine (infolge SP) hat zahlreiche Vorteile, wie beispielsweise Verminderung des allergenen Potentials oder Veränderung in der Zusammensetzung oder Menge anderer Metabolite. Speicherproteine sind u.a. beschrieben in EP-A 0 591 530, WO 87/47731, WO 98/26064, EP-A 0 620 281; Kohno-Murase J et al. (1994) Plant Mol Biol 26(4): 1115-1124.

SP dienen zur Speicherung von Kohlenstoff, Stickstoff und Schwefel, die für das schnelle heterotrophe Wachstum bei Keimung von Samen oder Pollen benötigt werden. Sie haben meist keine enzymatische Aktivität. SP werden dabei nur im Embryo während der Samenentwicklung synthetisiert und akkumulieren dabei zum einen in Proteinspeichervakuolen (PSV) von unterschiedlich differenzierten Zellen im Embryo bzw. Endosperm.

"Speicherprotein" meint allgemein ein Protein, das mindestens eine der nachfolgenden wesentlichen Eigenschaften aufweist:

i) Speicherproteine werden im wesentlichen nur im Embryo während der Samenentwicklung exprimiert. "Im wesentlichen" bedeutet dabei, dass in dem besagten Stadium mindestens 50%, bevorzugt mindestens 70%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% der Gesamtexpression über die Lebensdauer einer Pflanze hinweg stattfindet.

## 29

ii) Speicherproteine werden während der Keimung des Samens wieder abgebaut. Dabei beträgt der Abbau während der Keimung mindestens 20%, bevorzugt mindestens 50%, ganz besonders bevorzugt mindestens 80%.

5

iii) Speicherproteine machen einen wesentlichen Anteil am Gesamtproteingehalt des nicht-keimenden Samens aus. Bevorzugt macht das Speicherprotein in dem nicht-keimenden Samen der Wildtyp-Pflanze mehr als 5 Gew.% des Gesamtproteins aus, besonders bevorzugt mindestens 10 Gew.%, ganz besonders bevorzugt mindestens 20 Gew.%, am meisten bevorzugt mindestens 30 Gew.-%.

10

Bevorzugt weisen Speicherproteine 2 oder alle der oben genannten wesentlichen Eigenschaften i), ii) oder iii) auf.

15

Speicherproteine können in Untergruppen entsprechend weiterer charakteristischer Eigenschaften, wie beispielsweise ihrem Sedimentationskoeffizienten oder der Löslichkeit in verschiedenen Lösungen (Wasser, Salzlösung, Alkohol) aufgeteilt werden. Die Bestimmung des Sedimentationskoeffizienten kann in der dem Fachmann vertrauten Weise mittels Ultrazentrifugation durchgeführt werden (z.B. beschrieben bei Correia JJ (2000) Methods in Enzymology 321:81-100).

20

25

Insgesamt können vier grosse Genfamilien für Speicherproteine aufgrund ihrer Sequenzen zugeordnet werden: 2S-Albumine (Napin-ähnlich), 7S-Globuline (Phaseolin-ähnlich), 11S/12S-Globuline (Legumin-/Cruciferin-ähnlich) und die Zein-Prolamine.

30

2S Albumine sind weit verbreitet in Samen von Dikotyledonen, einschliesslich wichtiger kommerzieller Pflanzenfamilien wie Fabaceae (z.B. Sojabohne), Brassicaceae (z.B. Raps), Euphorbiaceae (z.B. Rizinus) oder Asteraceae (z.B. Sonnenblume). 2S Albumine sind kompakte globuläre Proteine mit konservierten Cysteinresten, die oft Heterodimere bilden.

35

7S-Globuline liegen in trimerer Form vor und enthalten keine Cysteinreste. Nach ihrer Synthese werden sie wie die 2S-Albumine in kleinere Fragmente gespalten und glykosyliert. Trotz Unterschiede in der Polypeptidgrösse sind die verschiedenen 7S-Globuline hoch konserviert und gehen vermutlich wie die 2S-Albumine auf einen gemeinsamen Vorläuferprotein zurück. Die 7S-Globuline sind nur in geringen Mengen in Monokotyledonen vorhanden. In Dikotyledonen ist ihr Anteil immer kleiner verglichen mit den 11S/12S-Globulinen.

40

45

## 30

11S/12S-Globuline stellen neben den 2S-Albuminen die Hauptfraktion der Speicherproteine in Dikotyledonen. Die hohe Ähnlichkeit der verschiedenen 11S-Globuline aus verschiedenen Pflanzengattungen lassen wiederum auf ein gemeinsames Vorläuferprotein in der Evolution schließen.

Bevorzugt ist das Speicherprotein ausgewählt aus den Klassen der 2S-Albumine (Napin-ähnlich), 7S-Globuline (Phaseolin-ähnlich), 11S/12S-Globuline (Legumin-/Cruciferin-ähnlich) oder Zein-Prolamine.

Besonders bevorzugte 2S-Albumine umfassen

- i) 2S-Albumine aus Arabidopsis, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 2, 4, 6 oder 8, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 1, 3, 5 oder 7 kodierten Proteine,
- ii) 2S-Albumine aus Arten der Gattung Brassica, wie beispielsweise Brassica napus, Brassica nigra, Brassica juncea, Brassica oleracea oder Sinapis alba, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 32, 34, 36, 38, 40, 46 oder 48, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 31, 33, 35, 37, 39, 45 oder 47 kodierten Proteine,
- iii) 2S-Albumine aus Soja, ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 42 oder 44, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 41 oder 43 kodierten Proteine,
- iv) 2S-Albumine aus Sonnenblume (Helianthus annuus), ganz besonders bevorzugt die 2S-Albumine mit der SEQ ID NO: 50 oder 52, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 49 oder 51 kodierten Proteine,

sowie die entsprechenden Homologen und funktionellen Äquivalente zu i) oder ii) oder iii) oder iv) aus identischen oder anderen Pflanzenarten, insbesondere Raps, Sonnenblume, Lein, Sesam, Färberdistel, Ölbaum, Soja oder verschiedene Nussarten. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 2S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Wasser aus.

## 31

- Funktionelle Äquivalente der 2S-Albumine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60%, bevorzugt mindestens 80%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% zu einer der Proteinsequenzen mit der SEQ ID NO: 2, 4, 6, 8, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50 oder 52 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 2S-Speicherproteins auf.
- Besonders bevorzugte 7S-Globuline umfassen solche aus Arabidopsis oder Soja, ganz besonders bevorzugt die Proteine mit der SEQ ID NO: 94 oder 96, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 93 oder 95 kodierten Proteine. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 7S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Salzlösung aus. Als weitere charakteristische Eigenschaft können 7S-Globuline keine Cysteinreste enthalten.
- Funktionelle Äquivalente der 7S-Globuline haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60%, bevorzugt mindestens 80%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% zu einer der Proteinsequenzen mit der SEQ ID NO: 94 oder 96 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 7S-Speicherproteins auf.
- Besonders bevorzugte 11S/12S-Globuline umfassen bevorzugt 11S-Globuline aus Raps, Soja und Arabidopsis insbesondere
- i) 11S-Globuline aus Raps mit der SEQ ID NO: 10, 12, 14, 16 oder 18, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 9, 11, 13, 15 oder 17 kodierten Proteine,

ii) die 11S-Globuline aus Soja mit der SEQ ID NO: 20, 22, 24, 26 oder 28, am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 19, 21, 23, 25 oder 27 kodierten Proteine,

5

iii) die 11S-Globuline aus *Arabidopsis thaliana* mit der SEQ ID NO: 60, 62, 64, 66, 68 oder 70 am meisten bevorzugt die durch die Nukleinsäuren gemäß SEQ ID NO: 59, 61, 63, 65, 67 oder 69 kodierten Proteine,

10

sowie die entsprechenden Homologen und funktionellen Äquivalente aus anderen Pflanzenarten, insbesondere Raps, Sonnenblume, Lein, Sesam, Färberdistel, Ölbaum, Soja oder verschiedene Nussarten, wie beispielsweise das Sonnenblume 11S Speicherprotein (SEQ ID NO: 30), insbesondere das durch die Nukleinsäuresequenz gemäß SEQ ID NO: 29 kodierte Protein. Funktionelle Äquivalente zeichnen sich bevorzugt neben den oben genannten wesentlichen Eigenschaften durch charakteristische Eigenschaften wie einen 11S- oder 12S-Sedimentationskoeffizienten und/oder durch eine Löslichkeit in Salzlösung (PBS; phosphatgepufferte Salzlösung) und/oder eine schlechte Löslichkeit in Wasser aus.

15

20

25

30

35

Funktionelle Äquivalente der 11S- oder 12S Albumine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60%, bevorzugt mindestens 80%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% zu einer der Proteinsequenzen mit der SEQ ID NO: 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 60, 62, 64, 66, 68 oder 70 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines 11S- oder 12S-Speicherproteins auf.

40

45

Besonders bevorzugte Zein-Prolamine umfassen bevorzugt solche aus monokotyledonen Pflanzen, insbesondere Mais, Reis, Hafer, Gerste oder Weizen. Ganz besonders bevorzugt sind die Mais Zein-Prolamine beschrieben durch SEQ ID NO: 98, 100, 102 oder 104 - insbesondere die durch SEQ ID NO 97, 99, 101 oder 103 kodierten Protein -, das Reis Prolamin gemäß SEQ ID NO: 106 - insbesondere das durch SEQ ID NO 105 kodierte Protein -, das Hafer Prolamin gemäß SEQ ID NO: 108 - insbesondere das durch SEQ ID NO 107 kodierte Proteine-, das Gerste Prolamin gemäß SEQ ID NO: 110 und/oder 111 - insbesondere das durch SEQ ID

## 33

NO 109 kodierte Protein - und das das Weizen Prolamin gemäß SEQ ID NO: 113 - insbesondere das durch SEQ ID NO 112 kodierte Protein. Funktionelle Äquivalente zeichnen sich bevorzugt durch eine Löslichkeit in 70%iger ethanolischer Lösung und eine schlechte Löslichkeit in Wasser oder Salzlösung aus.

Funktionelle Äquivalente der Zein-Prolamine haben in einer weiteren bevorzugten Ausführungsform eine Homologie von mindestens 60%, bevorzugt mindestens 80%, ganz besonders bevorzugt mindestens 90%, am meisten bevorzugt mindestens 95% zu einer der Proteinsequenzen mit der SEQ ID NO: 98, 100, 102, 104, 106, 108, 110, 111 oder 113 wobei die Homologie sich bevorzugt über eine Länge von mindestens 30 Aminosäuren, bevorzugt mindestens 50 Aminosäuren besonders bevorzugt über 100 Aminosäuren, am meisten bevorzugt über die gesamte Länge der jeweiligen Proteine erstreckt, und weisen die gleichen wesentlichen Eigenschaften eines Speicherproteins und - bevorzugt- die charakteristischen Eigenschaften eines Zein-Prolamine auf.

Funktionelle Äquivalente meint insbesondere natürliche oder künstliche Mutationen der obengenannten Speicherproteine sowie homologe Polypeptide aus anderen Pflanzen, die die gleichen wesentlichen und - bevorzugt - charakteristischen Eigenschaften aufweisen. Bevorzugt sind homologe Polypeptide aus oben beschriebenen bevorzugten Pflanzen. Die zu den im Rahmen dieser Erfindung offenbarten Speicherproteinen homologen Sequenzen aus anderen Pflanzen - beispielsweise solchen deren genomische Sequenz ganz oder teilweise bekannt ist, wie beispielsweise aus *Arabidopsis thaliana*, *Brassica napus*, *Nicotiana tabacum* oder *Solanum tuberosum* - durch Homologievergleiche aus Datenbanken auffinden. können z.B. durch Datenbanksuche oder Durchmustern von Gen-Banken - unter Verwendung der beispielhaft aufgeführten Speicherprotein-Sequenzen als Suchsequenz bzw. Sonde - leicht aufgefunden werden.

Mutationen umfassen Substitutionen, Additionen, Deletionen, Inversion oder Insertionen eines oder mehrerer Aminosäurereste.

Ein weiterer Gegenstand der Erfindung umfasst ein zumindest teilweise doppelsträngiges Ribonukleinsäuremolekül, wobei das doppelsträngige Ribonukleinsäuremolekül umfasst

i) einen "sense"-RNA-Strang umfassend mindestens zwei Ribonukleotidsequenzabschnitte, wobei jeweils mindestens einer dieser Ribonukleotidsequenzabschnitte im wesentlichen

## 34

identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 59, 61, 63, 65, 67, 69, 71, 93, 95, 97, 99, 101, 103, 105, 107, 109 oder 112 oder eines funktionellen Äquivalentes derselben, wobei jedoch nicht alle Ribonukleotidsequenzabschnitte zu dem "sense"-RNA-Transkript eines einzigen einer Speicherprotein-Nukleinsäuresequenz identisch sind, und

ii) einen "antisense"-RNA-Strang, der zu dem RNA-sense-Strang unter i) im wesentlichen komplementären ist.

Bevorzugt haben zumindest zwei der Speicherprotein-Nukleinsäuresequenzen, zu deren "sense"-RNA-Transkript die besagten Ribonukleotidsequenzabschnitte im wesentlichen identisch sind, untereinander eine Homologie von unter 90%, bevorzugt unter 80%, ganz besonders bevorzugt unter 60% am meisten bevorzugt unter 50% über die gesamte Länge ihrer kodierenden Nukleotidsequenz.

In einer weiteren bevorzugten Ausführungsform enthält die dsRNA mehrere Sequenzabschnitte, die eine gleichzeitige Suppression mehrerer Speicherproteine, bevorzugt von Speicherproteinen aus verschiedenen Klassen - wie beispielsweise einem 2S-Albumin, 7S-Globuline, 11S/12S-Globulin oder die Zein-Prolamine - bewirken.

Am meisten bevorzugt sind doppelsträngige RNA Moleküle beschrieben durch die Ribonukleinsäuresequenz gemäß SEQ ID NO: 84, 86 oder 88. Diese werden bevorzugt kodiert durch Nukleotidsequenzen entsprechend SEQ ID NO: 83, 85 oder 87.

5. Erreichen einer Resistenz gegen pflanzliche Pathogene

Eine Resistenz gegen pflanzliche Pathogene wie Arachniden, Pilze, Insekten, Nematoden, Protozoen, Viren, Bakterien und Krankheiten kann erreicht werden durch Verminderung der Genexpression von Genen, die für das Wachstum, Überleben, bestimmte Entwicklungsstufen (beispielsweise Verpuppung) oder die Vermehrung eines bestimmten Pathogens essentiell sind. Eine entsprechende Verminderung kann eine vollständige Inhibition vorgenannter Schritte aber auch eine Verzögerung derselben bewirken. Dies können pflanzliche Gene sein, die dem Pathogen beispielsweise das Eindringen ermöglichen, können aber auch pathogen-eigene Gene sein. Bevorzugt ist die dsRNA

## 35

gg. Gene des Pathogens gerichtet. Als anti-Pathogenes Agens kann dabei die dsRNA selber, jedoch auch die Expressionssysteme, Expressionskassetten oder transgenen Organismen wirken. Pflanzen können beispielsweise mit geeigneten Formulierungen vorgenannter Agentien behandelt, beispielsweise besprüht oder estäubt werden. Die Pflanzen selber können jedoch in Form eines transgenen Organismus die Agentien beinhalten und diese - beispielsweise in Form eines Fraßgiftes - an die Pathogene weitergeben. Verschiedene essentielle Gene diverser Pathogene sind dem Fachmann bekannt (z.B. für Nematodenresistenz WO 93/10251, WO 94/17194).

Am meisten bevorzugt als Pathogen sind Pilzpathogene wie *Phytophthora infestans*, *Fusarium nivale*, *Fusarium graminearum*, *Fusarium culmorum*, *Fusarium oxysporum*, *Blumeria graminis*, *Magnaporthe grisea*, *Sclerotinia sclerotium*, *Septoria nodorum*, *Septoria tritici*, *Alternaria brassicae*, *Phoma lingam*, bakterielle Pathogene wie *Corynebacterium sepedonicum*, *Erwinia carotovora*, *Erwinia amylovora*, *Streptomyces scabies*, *Pseudomonas syringae* pv. *tabaci*, *Pseudomonas syringae* pv. *phaseolicola*, *Pseudomonas syringae* pv. *tomato*, *Xanthomonas campestris* pv. *malvacearum* und *Xanthomonas campestris* pv. *oryzae*, und Nematoden wie *Globodera rostochiensis*, *G. pallida*, *Heterodera schachtii*, *Heterodera avenae*, *Ditylenchus dipsaci*, *Anguina tritici* und *Meloidogyne hapla*.

Eine Virusresistenz kann beispielsweise durch Verminderung der Expression eines viralen Hüllproteins, einer viralen Replikase, einer viralen Protease etc. erreicht werden. Zahlreiche Pflanzenviren und entsprechende Zielgene sind dem Fachmann bekannt.

#### 6. Verhinderung von Halmbruch

Eine verminderte Anfälligkeit gegen Halmbruch kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen des Kohlenhydratstoffwechsels (s.o.). Vorteilhafte Gene sind beschrieben (u.a. WO 97/13865) und umfassen gewebespezifische Polygalacturonasen oder Cellulasen.

#### 7. Verzögerung der Fruchtreifung

Eine verzögerte Fruchtreifung kann beispielsweise erreicht werden durch Verminderung der Genexpression von Genen ausgewählt aus der Gruppe bestehend aus Polygalacturonasen, Pectinesterasen,  $\beta$ -(1-4)glucanasen (Cellulasen),  $\beta$ -Galactanasen ( $\beta$ -Galactosidasen), oder Gene der Ethylenbiosynthese wie



## 36

- 1-Aminocyclopropan-1-carboxylatsynthase, Gene der Carotinoidbiosynthese wie z.B. Gene der Prephytoen- oder Phytoenbiosynthese beispielsweise Phytoendesaturasen. Weitere vorteilhafte Gene sind beispielsweise in WO 91/16440, WO 91/05865, WO 5 91/16426, WO 92/17596, WO 93/07275 oder WO 92/04456.
8. Erzielen einer männlichen Sterilität ("male sterility"). Entsprechende Zielgene sind beschrieben u.a. in WO 94/29465, WO89/10396, WO 92/18625.
- 10 9. Verminderung unerwünschter oder toxischer Pflanzeninhaltsstoffe wie z.B. Glucosinolaten. Entsprechende Zielgene sind beschrieben (u.a. in WO 97/16559).
- 15 10. Verzögerung von Alterserscheinungen. Entsprechende Zielgene sind u.a. Cinnamoyl-CoA:NADPH-Reduktasen oder Cinnamoylalkoholdehydrogenasen. Weitere Zielgene sind beschrieben (u.a. in WO 95/07993).
- 20 11. Modifikation der Lignifikation und/oder des Ligningehaltes vor allem in Baumarten. Entsprechende Zielgene sind beschrieben u.a. in WO 93/05159, WO 93/05160.
- 25 12. Modifikation des Faseranteils in Nahrungsmitteln vorzugsweise in Samen durch Verminderung der Expression der Coffeinsäure-O-methyltransferase oder der Cinnamoylalkoholdehydrogenase.
- 30 13. Modifikation der Faserqualität in Baumwolle. Entsprechende Zielgene sind beschrieben u.a. in US 5,597,718.
- 35 14. Verminderung der Stoßanfälligkeit von beispielsweise Kartoffeln durch Verminderung beispielsweise der Polyphenoloxidase (WO 94/03607) etc.
- 40 15. Steigerung der Vitamin E Biosynthese beispielsweise durch Verminderung der Expression von Genen aus dem Homogentisatabauweg wie z.B. der Homogentisat-1,2-dioxygenase (HGD; EC-Nr.: 1.13.11.5), der Maleylacetoacetatisomerase (MAAI; EC-Nr.: 5.2.1.2.) oder der Fumarylacetoacetathydrolase (FAAH; EC-Nr.: 3.7.1.2).

Ein weiterer Gegenstand der Erfindung umfasst ein zumindest teilweise doppelsträngiges Ribonukleinsäuremolekül, wobei das doppelsträngige Ribonukleinsäuremolekül umfasst

45

## 37

- i) einen "sense"-RNA-Strang umfassend mindestens zwei Ribonukleotidsequenzabschnitte, wobei jeweils mindestens einer dieser Ribonukleotidsequenzabschnitte im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines Gens aus dem Homogentisatabbauweg gemäß SEQ ID NO: 115, 116, 118 oder 120 oder eines funktionellen Äquivalentes derselben, wobei jedoch nicht alle Ribonukleotidsequenzabschnitte zu dem "sense"-RNA-Transkript eines einzigen einer Speicherprotein-Nukleinsäuresequenz identisch sind, und
- ii) einen "antisense"-RNA-Strang, der zu dem RNA-sense-Strang unter i) im wesentlichen komplementären ist.

16. Verminderung des Nikotingehaltes beispielsweise in Tabak durch verminderte Expression beispielsweise der N-Methylputrescinoxidase und der Putrescin-N-methyltransferase.
17. Verminderung des Coffeingehaltes in der Kaffeebohne (*Coffea arabica*) durch Verminderung der Genexpression von Genen der Coffeinbiosynthese wie 7-Methylxanthine-3-methyltransferase.
18. Verminderung des Theophyllin-Gehaltes im Tee (*Camellia sinensis*) durch Verminderung der Genexpression von Genen der Theophyllin-Biosynthese wie beispielsweise 1-Methylxanthin-3-methyltransferase.
19. Erhöhung des Methioningehaltes durch Verminderung der Threoninbiosynthese, beispielsweise durch Verminderung der Expression der Threoninsynthase (Zeh M et al. (2001) Plant Physiol 127(3):792-802).

Weitere Beispiele für vorteilhafte Gene sind zum Beispiel genannt bei Dunwell JM, Transgenic approaches to crop improvement, J Exp Bot. 2000;51 Spec No; Seite 487-96.

Jede der oben genannten Anwendungen kann als solche isoliert angewendet werden. Natürlich können auch mehr als eine der oben genannten Ansätze gleichzeitig angewendet werden. Dabei wird bei allen Anwendungen die Expression von mindestens zwei unterschiedlichen Zielgenen, wie oben definiert, vermindert. Diese Zielgene können dabei aus einer einzigen für eine Anwendung bevorzugten Gruppe von Genen stammen oder aber auch unterschiedlichen Anwendungsgruppen zugeordnet sein.

## 38

Zur Anwendung der erfindungsgemäßen Verfahren stehen dem Fachmann geläufige Werkzeuge, wie Expressionsvektoren mit für Pflanzen geeigneten Promotoren, sowie Verfahren zur Transformation und Regeneration von Pflanzen zur Verfügung. Pflanzenspezifische

- 5 Promotoren meint grundsätzlich jeden Promotor, der die Expression von Genen, insbesondere Fremdgenen, in Pflanzen oder Pflanzenteilen, -zellen, -geweben, -kulturen steuern kann. Dabei kann die Expression beispielsweise konstitutiv, induzierbar oder entwicklungsabhängig sein. Bevorzugt sind:

10

a) Konstitutive Promotoren

- "Konstitutive" Promotoren meint solche Promotoren, die eine Expression in zahlreichen, bevorzugt allen, Geweben über einen  
15 größeren Zeitraum der Pflanzenentwicklung, bevorzugt zu allen Zeitpunkten der Pflanzenentwicklung, gewährleisten (Benfey et al. (1989) EMBO J 8:2195-2202). Vorzugsweise verwendet man insbesondere einen pflanzlichen Promotor oder einen Promotor, der einem Pflanzenvirus entstammt. Insbesondere bevorzugt ist der  
20 Promotor des 35S-Transkriptes des CaMV Blumenkohlmosaikvirus (Franck et al. (1980) Cell 21:285-294; Odell et al. (1985) Nature 313:810-812; Shewmaker et al. (1985) Virology 140:281-288; Gardner et al. (1986) Plant Mol Biol 6:221-228) oder der 19S CaMV Promotor (US 5,352,605; WO 84/02913; Benfey et al. (1989)  
25 EMBO J 8:2195-2202). Ein weiterer geeigneter konstitutiver Promotor ist der "Rubisco small subunit (SSU)"-Promotor (US 4,962,028), der LeguminB-Promotor (GenBank Acc.-Nr. X03677), der Promotor der Nopalinsynthase aus Agrobacterium, der TR-Doppelpromotor, der OCS (Octopin Synthase) Promotor aus Agrobacterium, der Ubiquitin Promotor (Holtorf S et al. (1995) Plant Mol  
30 Biol 29:637-649), den Ubiquitin 1 Promotor (Christensen et al. (1992) Plant Mol Biol 18:675-689; Bruce et al. (1989) Proc Natl Acad Sci USA 86:9692-9696), den Smas Promotor, den Cinnamylalkoholdehydrogenase-Promotor (US 5,683,439), die Promotoren der  
35 vakuolärer ATPase Untereinheiten oder der Promotor eines prolinreichen Proteins aus Weizen (WO 91/13991), sowie weitere Promotoren von Genen, deren konstitutive Expression in Pflanzen dem Fachmann bekannt ist.

40 b) Gewebespezifische Promotoren

Bevorzugt sind ferner Promotoren mit Spezifitäten für die Ant-  
heren, Ovarien, Blüten, Blätter, Stengel, Wurzeln und Samen.

- 45 Samenspezifische Promotoren wie zum Beispiel der Promotor des Phaseolins (US 5,504,200; Bustos MM et al. (1989) Plant Cell 1(9):839-53), des 2S Albumingens (Joseffson LG et al. (1987) J

## 39

- Biol Chem 262:12196-12201), des Legumins (Shirsat A et al. (1989) Mol Gen Genet 215(2): 326-331), des USP (unknown seed protein; Bäumlein H et al. (1991) Mol Gen Genet 225(3):459-67), des Napin Gens (US 5,608,152; Stalberg K et al. (1996) L Planta 199:515-519), des Saccharosebindepoteins (WO 00/26388) oder der Legumin B4-Promotor (LeB4; Bäumlein H et al. (1991) Mol Gen Genet 225: 121-128; Bäumlein et al. (1992) Plant Journal 2(2):233-9; Fiedler U et al. (1995) Biotechnology (NY) 13(10):1090f), der Oleosin-Promoter aus Arabidopsis (WO 98/45461), der Bce4-Promoter aus Brassica (WO 91/13980). Weitere geeignete samenspezifische Promotoren sind die der Gene kodierend für das "High Molecular Weight Glutenin" (HMWG), Gliadin, Verzweigungsenzym, ADP Glucose Pyrophosphatase (AG-Pase) oder die Stärkesynthase. Bevorzugt sind ferner Promotoren, die eine samenspezifische Expression in Monokotyledonen wie Mais, Gerste, Weizen, Roggen, Reis etc. erlauben. Vorteilhaft eingesetzt werden können der Promoter des lpt2 oder lpt1-Gen (WO 95/15389, WO 95/23230) oder die Promotoren beschrieben in WO 99/16890 (Promotoren des Hordein-Gens, des Glutelin-Gens, des Oryzin-Gens, des Prolamin-Gens, des Gliadin-Gens, des Glutelin-Gens, des Zein-Gens, des Kasirin-Gens oder des Secalin-Gens). Weitere samenspezifische Promotoren sind beschrieben in WO89/03887.
- 25 Knollen-, Speicherwurzel- oder Wurzel-spezifische Promotoren wie beispielsweise der Patatin Promotor Klasse I (B33), der Promotor des Cathepsin D Inhibitors aus Kartoffel.
- 30 Blattspezifische Promotoren wie Promotor der cytosolischen FBPase aus Kartoffel (WO 97/05900), der SSU Promotor (small subunit) der Rubisco (Ribulose-1,5-bisphosphatcarboxylase) oder der ST-LSI Promotor aus Kartoffel (Stockhaus et al. (1989) EMBO J 8:2445-2451).
- 35 Blütenspezifische Promotoren wie beispielsweise der Phytoen Synthase Promotor (WO 92/16635) oder der Promotor des P-rr Gens (WO 98/22593).
- 40 - Antheren-spezifische Promotoren wie den 5126-Promotor (US 5,689,049, US 5,689,051), den glob-1 Promotor und den  $\gamma$ -Zein Promotor.

## c) Chemisch induzierbare Promotoren

- 45 Die Expressionskassetten können auch einen chemisch induzierbaren Promotor enthalten (Übersichtsartikel: Gatz et al. (1997) Annu Rev Plant Physiol Plant Mol Biol 48:89-108), durch den die

## 40

Expression des exogenen Gens in der Pflanze zu einem bestimmten Zeitpunkt gesteuert werden kann. Derartige Promotoren, wie z.B. der PRP1 Promotor (Ward et al. (1993) Plant Mol Biol 22:361-366), durch Salicylsäure induzierbarer Promotor (WO 5 95/19443), ein durch Benzolsulfonamid-induzierbarer Promotor (EP 0 388 186), ein durch Tetrazyklin-induzierbarer Promotor (Gatz et al. (1992) Plant J 2:397-404), ein durch Abscisinsäure induzierbarer Promotor (EP 0 335 528) bzw. ein durch Ethanol- oder Cyclohexanon-induzierbarer Promotor (WO 93/21334) können 10 ebenfalls verwendet werden.

## d) Stress- oder Pathogen-induzierbare Promotoren

Ferner sind Promotoren bevorzugt, die durch biotischen oder abiotischen Stress induziert werden wie beispielsweise der pathogen-induzierbare Promotor des PRP1-Gens (Ward et al. (1993) Plant Mol Biol 22:361-366), der hitzeinduzierbare hsp70- oder hsp80-Promoter aus Tomate (US 5,187,267), der kälteinduzierbare alpha-Amylase Promoter aus der Kartoffel (WO 96/12814), der 20 licht-induzierbare PPDK Promotor oder der verwundungsinduzierte pinII-Promoter (EP375091).

Pathogen-induzierbare Promotoren umfassen die von Genen, die infolge eines Pathogenbefalls induziert werden wie beispielsweise Gene von PR-Proteinen, SAR-Proteinen,  $\beta$ -1,3-Glucanase, Chitinase usw. (beispielsweise Redolfi et al. (1983) Neth J 25 Plant Pathol 89:245-254; Uknes, et al. (1992) The Plant Cell 4:645-656; Van Loon (1985) Plant Mol Viral 4:111-116; Marineau et al. (1987) Plant Mol Biol 9:335-342; Matton et al. (1987) 30 Molecular Plant-Microbe Interactions 2:325-342; Somssich et al. (1986) Proc Natl Acad Sci USA 83:2427-2430; Somssich et al. (1988) Mol Gen Genetics 2:93-98; Chen et al. (1996) Plant J 10:955-966; Zhang and Sing (1994) Proc Natl Acad Sci USA 91:2507-2511; Warner, et al. (1993) Plant J 3:191-201; Siebertz et al. (1989) Plant Cell 1:961-968(1989). 35

Umfasst sind auch verwundungs-induzierbare Promotoren wie der des pinII Gens (Ryan (1990) Ann Rev Phytopath 28:425-449; Duan et al. (1996) Nat Biotech 14:494-498), des wun1 und wun2-Gens 40 (US 5,428,148), des win1- und win2-Gens (Stanford et al. (1989) Mol Gen Genet 215:200-208), des Systemin (McGurl et al. (1992) Science 225:1570-1573), des WIP1-Gens (Rohmeier et al. (1993) Plant Mol Biol 22:783-792; Eckelkamp et al. (1993) FEBS Letters 323:73-76), des MPI-Gens (Corderok et al. (1994) The Plant J 45 6(2):141-150) und dergleichen.

## 41

## e) Entwicklungsabhängige Promotoren

- Weitere geeignete Promotoren sind beispielsweise fruchtreifungsspezifische Promotoren, wie beispielsweise der fruchtreifungsspezifische Promotor aus Tomate (WO 94/21794, EP 409 625). Entwicklungsabhängige Promotoren schließt zum Teil die Gewebespezifischen Promotoren ein, da die Ausbildung einzelner Gewebe naturgemäß entwicklungsabhängig erfolgt.

Besonders bevorzugt sind konstitutive sowie samenspezifische Promotoren.

Genetische Kontrollsequenzen umfassen auch weitere Promotoren, Promotorelemente oder Minimalpromotoren, die die expressionsteuernden Eigenschaften modifizieren können. So kann durch genetische Kontrollsequenzen zum Beispiel die gewebespezifische Expression zusätzlich abhängig von bestimmten Stressfaktoren erfolgen. Entsprechende Elemente sind zum Beispiel für Wasserstress, Abscissinsäure (Lam E und Chua NH, J Biol Chem 1991; 266(26): 17131-17135) und Hitzestress (Schoffl F et al., Molecular & General Genetics 217(2-3):246-53, 1989) beschrieben.

Genetische Kontrollsequenzen umfassen ferner auch die 5'-untranslatierte Regionen, Introns oder nichtkodierende 3'-Region von Genen wie beispielsweise das Actin-1 Intron, oder die Adh1-S Introns 1, 2 und 6 (allgemein: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994)). Es ist gezeigt worden, dass diese eine signifikante Funktionen bei der Regulation der Genexpression spielen können. So wurde gezeigt, dass 5'-untranslatierte Sequenzen die transiente Expression heterologer Gene verstärken können. Beispielhaft für Translationsverstärker sei die 5'-Leadersequenz aus dem Tabak-Mosaik-Virus zu nennen (Gallie et al. (1987) Nucl Acids Res 15:8693-8711) und dergleichen. Sie können ferner die Gewebsspezifität fördern (Rouster J et al. (1998) Plant J 15:435-440).

Als Kontrollsequenzen geeignete Polyadenylierungssignale sind pflanzliche Polyadenylierungssignale, vorzugsweise solche, die im wesentlichen T-DNA Polyadenylierungssignale aus *Agrobacterium tumefaciens*, insbesondere des Gens 3 der T-DNA (Octopin Synthase) des Ti-Plasmids pTiACHS entsprechen (Gielen et al. (1984) EMBO J 3:835 ff) oder funktionelle Äquivalente davon. Beispiele für besonders geeignete Terminatorsequenzen sind der OCS (Octopin-Synthase)-Terminator und der NOS (Nopalinsynthase)-Terminator.

Eine Expressionskassetten und die von ihr abgeleiteten Vektoren können weitere Funktionselemente enthalten. Der Begriff Funktionselement ist breit zu verstehen und meint all solche Elemente,

## 42

die einen Einfluss auf Herstellung, Vermehrung oder Funktion der erfindungsgemässen Expressionskassetten, Vektoren oder transgenen Organismen haben. Beispielfhaft aber nicht einschränkend seien zu nennen:

5

a) Selektionsmarker, die eine Resistenz gegen einen Metabolismusinhibitor wie 2-Desoxyglucose-6-phosphat (WO 98/45456), Antibiotika oder Biozide, bevorzugt Herbizide, wie zum Beispiel Kanamycin, G 418, Bleomycin, Hygromycin, oder Phosphinotricin etc. verleihen. Besonders bevorzugte Selektionsmarker sind solche die eine Resistenz gegen Herbizide verleihen. Beispielfhaft seien genannt: DNA Sequenzen, die für Phosphinotricinacetyltransferasen (PAT) kodieren und Glutaminsynthaseinhibitoren inaktivieren (bar und pat Gen), 5-Enolpyruvylshikimat-3-phosphatsynthasegene (EPSP Synthasegene), die eine Resistenz gegen Glyphosat® (N-(phosphonomethyl)glycin) verleihen, das für das Glyphosat® degradierende Enzyme kodierende gox Gen (Glyphosatoxidoreduktase), das deh Gen (kodierend für eine Dehalogenase, die Dalapon inaktiviert), Sulfonylurea- und Imidazolinon inaktivierende Acetolactatsynthasen sowie bxn Gene, die für Bromoxynil degradierende Nitrilaseenzyme kodieren, das aasa-Gen, das eine Resistenz gegen das Antibiotikum Apectinomycin verleiht, das Streptomycinphosphotransferase (SPT) Gen, das eine Resistenz gegen Streptomycin gewährt, das Neomycinphosphotransferase (NPTII) Gen, das eine Resistenz gegen Kanamycin oder Geneticidin verleiht, das Hygromycinphosphotransferase (HPT) Gen, das eine Resistenz gegen Hygromycin vermittelt, das Acetolactatsynthase Gen (ALS), das eine Resistenz gegen Sulfonylharnstoff-Herbizide verleiht (z.B. mutierte ALS-Varianten mit z.B. der S4 und/oder Hra Mutation).

30

b) Reportergene, die für leicht quantifizierbare Proteine kodieren und über Eigenfarbe oder Enzymaktivität eine Bewertung der Transformationseffizienz oder des Expressionsortes oder -zeitpunktes gewährleisten. Ganz besonders bevorzugt sind dabei Reporter-Proteine (Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) wie das "green fluorescence protein" (GFP) (Sheen et al. (1995) Plant Journal 8(5):777-784), die Chloramphenicoltransferase, eine Luziferase (Ow et al. (1986) Science 234:856-859), das Aequorin-Gen (Prasher et al. (1985) Biochem Biophys Res Commun 126(3):1259-1268), die  $\beta$ -Galactosidase, ganz besonders bevorzugt ist die  $\beta$ -Glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907).

45

## 43

- c) Replikationsursprünge, die eine Vermehrung der erfindungsge-  
mässen Expressionskassetten oder Vektoren in zum Beispiel  
E.coli gewährleisten. Beispielhaft seien genannt ORI (origin  
of DNA replication), der pBR322 ori oder der P15A ori (Sam-  
brook et al.: Molecular Cloning. A Laboratory Manual, 2nd ed.  
5 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY,  
1989).
- d) Elemente, die für eine Agrobakterium vermittelte Pflanzen-  
10 transformation erforderlich sind, wie zum Beispiel die rechte  
oder linke Begrenzung der T-DNA oder die vir-Region.

Zur Selektion erfolgreich homolog rekombinierter oder auch trans-  
formierter Zellen ist es in der Regel erforderlich, einen selek-  
15 tionierbaren Marker zusätzlich einzuführen, der den erfolgreich  
rekombinierten Zellen eine Resistenz gegen ein Biozid (zum Bei-  
spiel ein Herbizid), einen Metabolismusinhibitor wie 2-Desoxyglu-  
cose-6-phosphat (WO 98/45456) oder ein Antibiotikum verleiht. Der  
Selektionsmarker erlaubt die Selektion der transformierten Zellen  
20 von untransformierten (McCormick et al. (1986) Plant Cell Reports  
5:81-84).

Verschiedene Methoden und Vektoren zum Einschleusen von Genen in  
das Genom von Pflanzen sowie zur Regeneration von Pflanzen aus  
25 Pflanzengewebe oder Pflanzenzellen sind bekannt (Plant Molecular  
Biology and Biotechnology (CRC Press, Boca Raton, Florida), Kapi-  
tel 6/7, S. 71-119 (1993); White FF (1993) Vectors for Gene  
Transfer in Higher Plants; in: Transgenic Plants, Bd. 1, Engineer-  
ing and Utilization, Hrsgb.: Kung und Wu R, Academic Press,  
30 15-38; Jenes B et al. (1993) Techniques for Gene Transfer, in:  
Transgenic Plants, Bd. 1, Engineering and Utilization, Hrsgb.:  
Kung und R. Wu, Academic Press, S.128-143; Potrykus (1991) Annu  
Rev Plant Physiol Plant Molec Biol 42:205-225; Halford NG, Shewry  
PR (2000) Br Med Bull 56(1):62-73). Dazu zählen beispielhaft die  
35 oben erwähnten. Bei Pflanzen werden dabei die beschriebenen Me-  
thoden zur Transformation und Regeneration von Pflanzen aus  
Pflanzengewebe oder Pflanzenzellen zur transienten oder stabilen  
Transformation genutzt. Geeignete Methoden sind vor allem die  
Protoplastentransformation durch Polyethylenglykol-induzierte  
40 DNA-Aufnahme, die Liposomen vermittelte Transformation (wie z.B.  
in US 4,536,475 beschrieben), biolistische Verfahren mit der Gen-  
kanone ("particle bombardment" Methode; Fromm ME et al. (1990)  
Bio/Technology. 8(9):833-9; Gordon-Kamm et al. (1990) The Plant  
Cell 2:603), die Elektroporation, die Inkubation trockener Em-  
45 bryonen in DNA-haltiger Lösung und die Mikroinjektion. Im Falle  
dieser "direkten" Transformationsmethoden sind keine besonderen  
Anforderungen an das verwendete Plasmid gestellt. Einfache Plas-



mide wie die der pUC-Reihe, pBR322, M13mp Reihe, pACYC184 etc. können verwendet werden. Sollen vollständige Pflanzen aus den transformierten Zellen regeneriert werden, so ist er erforderlich, das sich auf dem Plasmid ein zusätzliches selektionierbares 5 Markergen befindet.

Neben diesen "direkten" Transformationstechniken kann eine Transformation auch durch bakterielle Infektion mittels *Agrobacterium* (z.B. EP 0 116 718), virale Infektion mittels viraler Vektoren 10 (EP 0 067 553; US 4,407,956; WO 95/34668; WO 93/03161) oder mittels Pollen (EP 0 270 356; WO 85/01856; US 4,684,611) durchgeführt werden.

Die für die *Agrobacterium*-Transformation meist verwendeten 15 Stämme *Agrobacterium tumefaciens* oder *Agrobacterium rhizogenes* eine auch durch bakterielle Infektion mittels enthalten ein Plasmid (Ti bzw. Ri Plasmid), das auf die Pflanze nach *Agrobacterium*-Infektion übertragen wird. Ein Teil dieses Plasmids, genannt T-DNA (transferred DNA), wird in das Genom der Pflanzenzelle integriert. Alternativ können durch *Agrobacterium* auch binäre vektoren 20 (Mini-Ti-Plasmide) auf Pflanzen übertragen und in deren Genom integriert werden. Die *Agrobacterium*-vermittelte Transformation ist am besten für dicotyledone, diploide Pflanzenzellen geeignet, wohingegen die direkten Transformationstechniken sich für jeden 25 Zelltyp eignen. Verfahren zur *Agrobacterium* vermittelten Transformation sind beispielsweise beschrieben bei Horsch RB et al. (1985) Science 225:1229f. Werden *Agrobacterien* verwendet, so ist die Expressionskassette in spezielle Plasmide zu integrieren, entweder in einen Zwischenvektor (englisch: shuttle or intermediate vector) oder einen binären Vektor. Wird ein Ti oder Ri 30 Plasmid zur Transformation verwendet werden soll, ist zumindest die rechte Begrenzung, meistens jedoch die rechte und die linke Begrenzung der Ti oder Ri Plasmid T-DNA als flankierende Region mit der einzuführenden Expressionskassette verbunden.

35 Für die *Agrobacterium* Transformation werden bevorzugt binäre Vektoren verwendet. Binäre Vektoren können sowohl in *E.coli* als auch in *Agrobacterium* replizieren. Sie enthalten in der Regel ein Selektionsmarkergen und einen Linker oder Polylinker flankiert von 40 der rechten und linken T-DNA Begrenzungssequenz. Sie können direkt in *Agrobacterium* transformiert werden (Holsters et al. (1978) Mol Gen Genet 163:181-187). Das Selektionsmarkergen erlaubt eine Selektion transformierter *Agrobakteria* und ist zum Beispiel das nptII Gen, das eine Resistenz gegen Kanamycin verleiht. Das in diesem Fall als Wirtsorganismus fungierende 45 *Agrobacterium* sollte bereits ein Plasmid mit der vir-Region enthalten. Diese ist für die Übertragung der T-DNA auf die pflanzliche

## 45

Zelle erforderlich. Ein so transformiertes Agrobacterium kann zur Transformation pflanzlicher Zellen verwendet werden. Die Verwendung von T-DNA zur Transformation pflanzlicher Zellen ist intensiv untersucht und beschrieben (EP 120 516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij KanTERS B.V., Alblas-  
5 sérdam, Chapter V; An et al. (1985) EMBO J 4:277-287). Verschiedene binäre Vektoren sind bekannt und teilweise kommerziell erhältlich wie zum Beispiel pBI101.2 oder pBIN19 (Clontech Laboratories, Inc. USA; Bevan et al. (1984) Nucl Acids Res 12:8711),  
10 pBinAR, pPZP200 oder pPTV.

Die mit einem solchen Vektor transformierten Agrobakterien können dann in bekannter Weise zur Transformation von Pflanzen, insbesondere von Kulturpflanzen, wie z.B. von Raps, verwendet werden, indem beispielsweise verwundete Blätter oder Blattstücke in  
15 einer Agrobakterienlösung gebadet und anschliessend in geeigneten Medien kultiviert werden. Die Transformation von Pflanzen durch Agrobakterien ist beschrieben (White FF, Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu,  
20 Academic Press, 1993, S. 15 - 38; Jenes B et al. (1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, herausgegeben von S.D. Kung und R. Wu, Academic Press, S.128-143; Potrykus (1991) Annu Rev Plant Physiol Plant  
25 Molec Biol 42:205- 225). Aus den transformierten Zellen der verwundeten Blätter bzw. Blattstücke können in bekannter Weise transgene Pflanzen regeneriert werden, die integriert die oben beschriebenen erfindungsgemässen Expressionssysteme enthalten.

30 Stabil transformierte Zellen d.h. solche, die die eingeführte DNA integriert in die DNA der Wirtszelle enthalten, können von untransformierten selektioniert werden, wenn ein selektionierbarer Marker Bestandteil der eingeführten DNA ist. Als Marker kann beispielhaft jedes Gen fungieren, dass eine Resistenz gegen Antibio-  
35 tika oder Herbizide (wie Kanamycin, G418, Bleomycin, Hygromycin oder Phosphinotricin etc.) zu verleihen vermag (s.o.). Transformierte Zellen, die ein solches Markergen exprimieren, sind in der Lage, in der Gegenwart von Konzentrationen eines entsprechenden Antibiotikums oder Herbizides zu überleben, die einen untransfor-  
40 mierten Wildtyp abtöten. Beispiel sind oben genannt und umfassen bevorzugt das bar Gen, dass Resistenz gegen das Herbizid Phosphinotricin verleiht. (Rathore KS et al. (1993) Plant Mol Biol 21(5):871-884), das nptII Gen, dass Resistenz gegen Kanamycin verleiht, das hpt Gen, das Resistenz gegen Hygromycin verleiht,  
45 oder das EPSP-Gen, das Resistenz gegen das Herbizid Glyphosat verleiht. Der Selektionsmarker erlaubt die Selektion von transformierten Zellen von untransformierten (McCormick et al. (1986)

Plant Cell Reports 5:81-84). Die erhaltenen Pflanzen können in üblicher Weise gezüchtet und gekreuzt werden. Zwei oder mehr Generationen sollten vorzugsweise kultiviert werden, um sicherzustellen, dass die genomische Integration stabil und vererblich ist.

Sobald eine transformierte Pflanzenzelle hergestellt wurde, kann eine vollständige Pflanze unter Verwendung von dem Fachmann bekannten Verfahren erhalten werden. Hierbei geht man beispielhaft von Kalluskulturen aus. Aus diesen noch undifferenzierten Zellmassen kann die Bildung von Spross und Wurzel in bekannter Weise induziert werden. Die erhaltenen Sprösslinge können ausgepflanzt und gezüchtet werden. Entsprechende Verfahren sind beschrieben (Fennell et al. (1992) Plant Cell Rep. 11: 567-570; Stoeger et al (1995) Plant Cell Rep. 14:273-278; Jahne et al. (1994) Theor Appl Genet 89:525-533).

Die Wirksamkeit der Expression der transgen exprimierten Nukleinsäuren kann beispielsweise *in vitro* durch Sprossmeristemvermehrung unter Verwendung einer der oben beschriebenen Selektionsmethoden ermittelt werden. Zudem kann eine in Art und Höhe veränderte Expression eines Zielgens und die Auswirkung auf den Phänotyp der Pflanze an Testpflanzen in Gewächshausversuchen getestet werden.

25

## II. Medizinische Anwendungen

Die erfindungsgemäß bereitgestellten dsRNA, Expressionssysteme oder Organismen eignen sich zur Herstellung von Arzneimitteln zur Behandlung von menschlichen und tierischen Erkrankungen. Für eine effiziente Therapie ist es oft unzureichend nur ein einzelnes Zielgen zu vermindern. Das erfindungsgemäße Verfahren eignet sich insbesondere zur Behandlung von

- Pathogenbefall, wie beispielsweise virale oder bakterielle Erkrankungen. In diesen Fällen führen Ansätze, die lediglich gegen ein molekulares Ziel gerichtet sind, oft zu der Ausbildung von Resistenzen. Eine Kombinationstherapie, die mehrere Ziele abdeckt, ist jedoch kompliziert zu koordinieren und v.a. nur sehr aufwendig in klinischen Experimenten zu evaluieren. Das erfindungsgemäße Verfahren ermöglicht hier eine vorteilhafte Alternative. Die inhibitorische dsRNA kann dabei in der dem Fachmann geläufigen Weise appliziert werden. dsRNA verfügt über eine erstaunliche Stabilität und effiziente Wirkung und kann beispielsweise durch Verfütterung entsprechender dsRNA exprimierenden Bakterien appliziert werden. Das Verfahren

## 47

eignet sich insbesondere zur Behandlung von viralen Infektionen z.B. mit dem "human immunodeficiency virus" (HIV), indem gleichzeitig die Expression von mindestens zwei viralen Genen vermindert wird, beispielsweise bei HIV von Genen wie gp41, die für den Zelleintritt verantwortlich sind, und der viralen Replikase oder reversen Transkriptase.

- Behandlung von Krebs (beispielsweise solider Tumore und/oder Leukämien). Zahlreiche potentialle Zielgene sind hier dem Fachmann bekannt (z.B. Oncogene wie ABL1, BCL1, BCL2, BCL6, CBFA2, CBL, CSF1R, ERBA, ERBB, EBRB2, FGR, FOS, FYN, HRAS, JUN, LCK, LYN, MYB, MYC, NRAS, RET oder SRC; Tumorsuppressorgene wie BRCA1 oder BRCA2; Adhäsionsmoleküle; Cyclinekinasen und deren Inhibitoren).

Weitere potentiell mit dem erfindungsgemäßen Verfahren behandelbare Erkrankungen und die entsprechenden Zielgene sind dem Fachmann ohne weiteres zugänglich und umfassen beispielsweise Erkrankungen des Herz/Kreislaufsystems wie Bluthochdruck, Erkrankungen des zentralen oder peripheren Nervensystems wie Alzheimer, Parkinson oder multiple Sklerose usw. Auch ist es durch das erfindungsgemäße Verfahren möglich, mehr als eine Erkrankung parallel zu behandeln, wie beispielsweise ein Herzkreislauferkrankung und eine Erkrankung des zentralen Nervensystems, was durch klassische Ansätze nicht möglich ist. Derartige Ansätze sind v.a. bei multiplen Erkrankungen wie sie oft im fortgeschrittenen Alter auftreten vorteilhaft. Beispielfhaft sei die parallele Behandlung von Bluthochdruck und z.B. Alzheimer oder seniler Demenz zu nennen. Dabei kann dieser Anwendungen als solche isoliert angewendet werden. Natürlich können auch mehr als eine der oben genannten Ansätze gleichzeitig angewendet werden. Dabei wird bei allen Anwendungen die Expression von mindestens zwei unterschiedlichen Zielgenen vermindert. Diese Zielgene können dabei aus der für eine Anwendung bevorzugten Gruppe von Genen stammen oder aber auch unterschiedlichen Anwendungsgruppen zugeordnet sein.

### III. Biotechnologische Anwendungen

40

Das erfindungsgemäße Verfahren läßt sich vorteilhaft in biotechnologischen Verfahren anwenden. Beispielfhaft jedoch nicht einschränkend sei zu nennen die Optimierung von Stoffwechselwegen in fermentativ genutzten Hefen, Pilzen oder anderen eukaryotischen Mikroorganismen oder Zellen zur Herstellung von Feinchemikalien wie Aminosäuren (z.B. Lysin oder Methionin), Vitaminen (wie Vitamin B2, Vitamin C, Vitamin E), Carotinoi-

45

den, Ölen und Fetten, polyungesättigten Fettsäuren, Biotin usw. Dabei kann dieser Anwendungen als solche isoliert angewendet werden. Natürlich können auch mehr als eine der oben genannten Ansätze gleichzeitig angewendet werden. Dabei wird bei allen Anwendungen die Expression von mindestens zwei unterschiedlichen Zielgenen vermindert. Diese Zielgene können dabei aus der für eine Anwendung bevorzugten Gruppe von Genen stammen oder aber auch unterschiedlichen Anwendungsgruppen zugeordnet sein.

10

Als Vektoren zur Expression in E.coli sind bevorzugt pQE70, pQE60 und pQE-9 (QIAGEN, Inc.); pBluescript Vektoren, Phagescript Vektoren, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene Cloning Systems, Inc.); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia Biotech, Inc.).

Bevorzugte Vektoren zur eukaryotischen Expression umfassen pWLNE0, pSV2CAT, pOG44, pXT1 und pSG (Stratagene Inc.); pSVK3, pBPV, pMSG und pSVL (Pharmacia Biotech, Inc.). Als induzierbare Vektoren seien pTet-tTak, pTet-Splice, pcDNA4/TO, pcDNA4/TO / LacZ, pcDNA6/TR, pcDNA4/TO/Myc-His /LacZ, pcDNA4/TO/Myc-His A, pcDNA4/TO/Myc-His B, pcDNA4/TO/Myc-His C, pVgRXR (Invitrogen, Inc.) oder die pMAM-Serie (Clontech, Inc.; GenBank Accession No.: U02443) zunennen. Diese stellen bereits das induzierbare regulatorische Kontrollelement beispielsweise für eine chemisch, induzierbare Expression eines DSBI-Enzyms zur Verfügung. In diese Vektoren kann die Nukleinsäuresequenz kodierend für ein DSBI-Enzym direkt insertiert werden.

Vektoren für die Expression in Hefe umfassen beispielhaft pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, PHIL-D2, PHIL-S1, pPIC3SK, pPIC9K, und PA0815 (Invitrogen, Inc.).

Vorteilhafte Kontrollsequenzen sind beispielsweise die gram-positiven Promotoren amy und SPO2, und die Hefe- oder Pilzpromotoren ADC1, MFa, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH.

Klonierungsvektoren und Techniken zur genetischen Manipulation von Ciliaten und Algen sind dem Fachmann bekannt (WO 98/01572; Falciatore et al. (1999) Marine Biotechnology 1(3):239-251; Dunahay et al. (1995) J Phycol 31:10004-1012).

Als Selektionsmarker können prinzipiell viele der auch für Pflanzen bevorzugten Selektionssysteme verwendet werden. Insbesondere bevorzugt sind für Säugerzelle die Neomycin (G418) Resistenz, die Hygromycin-Resistenz, die Zeocin-Resistenz oder die Puromycin-Re-

## 49

sistenz. Für Prokaryoten ist insbesondere die Ampicillin-Resistenz, die Kanamycin-Resistenz oder die Tetracyclin-resistent bevorzugt.

5 Prinzipiell sind für die Transformation tierischer Zelle oder von Hefezellen ähnliche Verfahren wie für die "direkte" Transformation von pflanzlichen Zellen anzuwenden. Insbesondere Verfahren wie die Calciumphosphat oder Liposomen vermittelte Transformation oder aber Elektroporation sind bevorzugt.

10

Ein weiterer Gegenstand der Erfindung betrifft die Verwendung der erfindungsgemässen, transgenen Organismen und der von ihnen abgeleitete Zellen, Zellkulturen, Teile - wie zum Beispiel bei transgenen pflanzlichen Organismen Wurzeln, Blätter etc.- , und trans-

15 genes Vermehrungsgut wie Saaten oder Früchte, zur Herstellung von Nahrungs- oder Futtermitteln, Pharmazeutika oder Feinchemikalien, wie beispielsweise Enzymen, Vitaminen, Aminosäuren, Zucker, Fettsäuren, natürliche und synthetische Geschmacks-, Aroma- und Farbstoffe. Besonders bevorzugt ist die Produktion von Triacylglyceriden, Lipiden, Ölen, Fettsäuren, Stärke, Tocopherolen und Tocotrienolen sowie Carotinoiden. Von Menschen und Tieren verzehrbare erfindungsgemässe, genetisch veränderte Pflanzen können auch beispielsweise direkt oder nach an sich bekannter Aufbereitung als Nahrungsmittel oder Futtermittel verwendet werden.

25

30

35

40

45

## Sequenzen

1. SEQ ID NO: 1  
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
5 unit 1 (GenBank Acc.-No.: M22032)
2. SEQ ID NO: 2  
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 1
- 10 3. SEQ ID NO: 3  
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
unit 3 (GenBank Acc.-No.: M22035)
4. SEQ ID NO: 4  
15 Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 3
5. SEQ ID NO: 5  
Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
unit 2 (GenBank Acc.-No.: M22034)  
20
6. SEQ ID NO: 6  
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 2
7. SEQ ID NO: 7  
25 Nukleinsäuresequenz kodierend für A.thaliana Albumin 2S sub-  
unit 4 (GenBank Acc.-No.: M22033)
8. SEQ ID NO: 8  
Proteinsequenz kodierend für A.thaliana Albumin 2S subunit 4  
30
9. SEQ ID NO: 9  
Nukleinsäuresequenz kodierend für B.napus Cruciferin Spei-  
cherprotein (GenBank Acc.-No.: X59294)
- 35 10. SEQ ID NO: 10  
Proteinsequenz kodierend für B.napus Cruciferin Speicherpro-  
tein
11. SEQ ID NO: 11  
40 Nukleinsäuresequenz kodierend für Brassica napus Cruciferin  
(GenBank Acc.-No.: X14555)
12. SEQ ID NO: 12  
Proteinsequenz kodierend für Brassica napus Cruciferin  
45

## 51

13. SEQ ID NO: 13  
Nukleinsäuresequenz kodierend für B.napus BnC2 Cruciferin  
Speicherprotein (GenBank Acc.-No.: X59295)
- 5 14. SEQ ID NO: 14  
Proteinsequenz kodierend für B.napus BnC2 Cruciferin Spei-  
cherprotein
15. SEQ ID NO: 15  
10 partielle Nukleinsäuresequenz kodierend für B.napus Crucife-  
rin cru4 subunit (GenBank Acc.-No.: X57848)
16. SEQ ID NO: 16  
partielle Proteinsequenz kodierend für B.napus Cruciferin  
15 cru4 subunit
17. SEQ ID NO: 17  
Nukleinsäuresequenz kodierend für B.napus cru1 Cruciferin  
subunit (GenBank Acc.-No.: X62120)  
20
18. SEQ ID NO: 18  
Proteinsequenz kodierend für B.napus cru1 Cruciferin subunit
19. SEQ ID NO: 19  
25 Nukleinsäuresequenz kodierend für Glycinin A-1a-B-x subunit  
aus des Sojabohne (GenBank Acc.-No.: M36686)
20. SEQ ID NO: 20  
Proteinsequenz kodierend für Glycinin A-1a-B-x subunit aus  
30 des Sojabohne
21. SEQ ID NO: 21  
Nukleinsäuresequenz kodierend für Sojabohne Glycinin subunit  
G2 (GenBank Acc.-No.: X15122)  
35
22. SEQ ID NO: 22  
Proteinsequenz kodierend für Sojabohne Glycinin subunit G2
23. SEQ ID NO: 23  
40 Nukleinsäuresequenz kodierend für Sojabohne A5A4B3 Glycinin  
subunits (GenBank Acc.-No.: X02626)
24. SEQ ID NO: 24  
Proteinsequenz kodierend für Sojabohne A5A4B3 Glycinin sub-  
45 units



## 52

25. SEQ ID NO: 25  
Nukleinsäuresequenz kodierend für Sojabohne (G.max) Glycinin  
Speicherprotein subunit A3-B4 (GenBank Acc.-No.: M10962)
- 5 26. SEQ ID NO: 26  
Proteinsequenz kodierend für Sojabohne (G.max) Glycinin Spei-  
cherprotein subunit A3-B4
- 10 27. SEQ ID NO: 27  
Nukleinsäuresequenz kodierend für Sojabohne Glycinin subunit  
G3 (GenBank Acc.-No.: X15123)
- 15 28. SEQ ID NO: 28  
Proteinsequenz kodierend für Sojabohne Glycinin subunit G3
29. SEQ ID NO: 29  
Nukleinsäuresequenz kodierend für Sonnenblume 11S Speicher-  
protein (G3-D1) (GenBank Acc.-No.: M28832)
- 20 30. SEQ ID NO: 30  
Proteinsequenz kodierend für Sonnenblume 11S Speicherprotein  
(G3-D1)
- 25 31. SEQ ID NO: 31  
Nukleinsäuresequenz kodierend für Raps (B.napus) Napin (Gen-  
Bank Acc.-No.: J02586)
- 30 32. SEQ ID NO: 32  
Proteinsequenz kodierend für Raps (B.napus) Napin
33. SEQ ID NO: 33  
Nukleinsäuresequenz kodierend für Brassica juncea 2S Spei-  
cherprotein (GenBank Acc.-No.: X65040)
- 35 34. SEQ ID NO: 34  
Proteinsequenz kodierend für Brassica juncea 2S Speicherpro-  
tein
- 40 35. SEQ ID NO: 35  
Nukleinsäuresequenz kodierend für Brassica oleracea 2S Spei-  
cherprotein (GenBank Acc.-No.: X65038)
- 45 36. SEQ ID NO: 36  
Proteinsequenz kodierend für Brassica oleracea 2S Speicher-  
protein

## 53

37. SEQ ID NO: 37  
Nukleinsäuresequenz kodierend für Brassica napus cv. Topas  
Napin (GenBank Acc.-No.: U04945)
- 5 38. SEQ ID NO: 38  
Proteinsequenz kodierend für Brassica napus cv. Topas Napin
39. SEQ ID NO: 39  
partielle Nukleinsäuresequenz kodierend für Sinapis alba sin1  
10 Speicherprotein (GenBank Acc.-No.: X91799)
40. SEQ ID NO: 40  
partielle Proteinsequenz kodierend für Sinapis alba sin1  
Speicherprotein
- 15 41. SEQ ID NO: 41  
Nukleinsäuresequenz kodierend für Sojabohne (Glycine max) na-  
pin-type 2S Albumin 1 (GenBank Acc.-No.: U71194)
- 20 42. SEQ ID NO: 42  
Proteinsequenz kodierend für Sojabohne (Glycine max) napin-  
type 2S Albumin 1
43. SEQ ID NO: 43  
25 Nukleinsäuresequenz kodierend für Sojabohne (Glycine max) 2S  
Albumin (GenBank Acc.-No.: AF005030)
44. SEQ ID NO: 44  
Proteinsequenz kodierend für Sojabohne (Glycine max) 2S Albu-  
30 min
45. SEQ ID NO: 45  
Nukleinsäuresequenz kodierend für Brassica nigra 2S Speicher-  
protein (GenBank Acc.-No.: X65039)
- 35 46. SEQ ID NO: 46  
Proteinsequenz kodierend für Brassica nigra 2S Speicherpro-  
tein
- 40 47. SEQ ID NO: 47  
Nukleinsäuresequenz kodierend für Sinapis alba sin5 Speicher-  
protein (GenBank Acc.-No.: X91798)
48. SEQ ID NO: 48  
45 Proteinsequenz kodierend für Sinapis alba sin5 Speicherpro-  
tein

## 54

49. SEQ ID NO: 49  
Nukleinsäuresequenz kodierend für Sonnenblume HaG5 2 S Albumin (GenBank Acc.-No.: X06410)
- 5 50. SEQ ID NO: 50  
proteinsequenz kodierend für Sonnenblume HaG5 2 S Albumin
51. SEQ ID NO: 51  
partielle Nukleinsäuresequenz kodierend für Sonnenblume (Helianthus annuus) 2S Albumin (GenBank Acc.-No.: X76101)
- 10
52. SEQ ID NO: 52  
partielle Proteinsequenz kodierend für Sonnenblume (Helianthus annuus) 2S Albumin
- 15
53. SEQ ID NO: 53  
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCru3 (Insert von Vektor pCR2.1-AtCRU3-RNAi)
- 20
54. SEQ ID NO: 54  
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCru3
- 25
55. SEQ ID NO: 55  
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCra1
- 30
56. SEQ ID NO: 56  
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 12S Speicherprotein AtCra1
- 35
57. SEQ ID NO: 57  
Nukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 2S Speicherprotein At2S2
- 40
58. SEQ ID NO: 58  
Ribonukleinsäuresequenz kodierend für dsRNA zur Suppression von Arabidopsis thaliana 2S Speicherprotein At2S2
- 45
59. SEQ ID NO: 59  
Nukleinsäuresequenz kodierend für Arabidopsis thaliana 12S Cruciferin Speicherprotein (ATCRU3; GenBank Acc.-No.: U66916)

## 55

60. SEQ ID NO: 60  
Proteinsequenz kodierend für *Arabidopsis thaliana* 12S Cruciferin Speicherprotein (ATCRU3)
- 5 61. SEQ ID NO: 61  
Nukleinsäuresequenz kodierend für *A.thaliana* 12S Speicherprotein (CRA1; GenBank Acc.-No.: M37247)
62. EQ ID NO: 62  
10 Proteinsequenz kodierend für *A.thaliana* 12S Speicherprotein (CRA1)
63. SEQ ID NO: 63  
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana* 12S Speicherprotein AT5g44120/MLN1\_4 (GenBank Acc.-No.: AY070730)  
15
64. SEQ ID NO: 64  
Proteinsequenz kodierend für *Arabidopsis thaliana* 12S Speicherprotein AT5g44120/MLN1\_4  
20
65. SEQ ID NO: 65  
Nukleinsäuresequenz kodierend für *Arabidopsis* 12S Speicherprotein (CRB; GenBank Acc.-No.: X14313; M37248)
- 25 66. SEQ ID NO: 66  
Proteinsequenz kodierend für *Arabidopsis* 12S Speicherprotein (CRB)
67. SEQ ID NO: 67  
30 Nukleinsäuresequenz kodierend für *Arabidopsis thaliana* putatives 12S Speicherprotein (aus GenBank Acc.-No.: AC003027)
68. SEQ ID NO: 68  
Proteinsequenz kodierend für *Arabidopsis thaliana* putatives Speicherprotein (Protein\_id="AAD10679.1")  
35
69. SEQ ID NO: 69  
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana* Cruciferin 12S Speicherprotein (Atlg03890) (GenBank Acc.-No.: AY065432)  
40
70. SEQ ID NO: 70  
Proteinsequenz kodierend für *Arabidopsis thaliana* Cruciferin 12S Speicherprotein (Atlg03890)  
45

71. SEQ ID NO: 71  
Nukleinsäuresequenz kodierend für *Arabidopsis thaliana* Prohibitin 1 (Atphb1) (GenBank Acc.-No.: U66594)
- 5 72. SEQ ID NO: 72  
Proteinsequenz kodierend für *Arabidopsis thaliana* Prohibitin 1 (Atphb1)
- 10 73. SEQ ID NO: 73 Oligonukleotidprimer OPN1
74. SEQ ID NO: 74 Oligonukleotidprimer OPN2
75. SEQ ID NO: 75 Oligonukleotidprimer OPN3
- 15 76. SEQ ID NO: 76 Oligonukleotidprimer OPN4
77. SEQ ID NO: 77 Oligonukleotidprimer OPN5
78. SEQ ID NO: 78 Oligonukleotidprimer OPN6
- 20 79. SEQ ID NO: 79 Oligonukleotidprimer OPN7
80. SEQ ID NO: 80 Oligonukleotidprimer OPN8
- 25 81. SEQ ID NO: 81 Oligonukleotidprimer OPN9
82. SEQ ID NO: 82 Oligonukleotidprimer OPN10
83. SEQ ID NO: 83
- 30 Nukleinsäuresequenz kodierend für sRNAi4-dsRNA zur Suppression mehrerer Speicherproteine
84. SEQ ID NO: 84
- 35 Ribonukleinsäuresequenz kodierend für sRNAi4-dsRNA zur Suppression mehrerer Speicherproteine
85. SEQ ID NO: 85
- 40 Nukleinsäuresequenz kodierend für sRNAi8-dsRNA zur Suppression mehrerer Speicherproteine
86. SEQ ID NO: 86
- 45 87. SEQ ID NO: 87 Oligonukleotidprimer OPN11
- Ribonukleinsäuresequenz kodierend für sRNAi8-dsRNA zur Suppression mehrerer Speicherproteine

## 57

88. SEQ ID NO: 88 Oligonukleotidprimer OP12
89. SEQ ID NO: 89 Oligonukleotidprimer OPN13
- 5 90. SEQ ID NO: 90 Oligonukleotidprimer OPN15
91. SEQ ID NO: 91 Oligonukleotidprimer OPN16
92. SEQ ID NO: 92 Oligonukleotidprimer OPN17
- 10 93. SEQ ID NO: 93  
Nukleinsäuresequenz kodierend für Arabidopsis thaliana "globulin-like protein" (GenBank Acc.-No.: NM\_119834)
- 15 94. SEQ ID NO: 94  
Proteinsequenz kodierend für Arabidopsis thaliana "globulin-like protein" (Protein\_id="NP\_195388.1")
95. SEQ ID NO: 95
- 20 96. Nukleinsäuresequenz kodierend für Glycine max 7S Samenglobulin (GenBank Acc.-No.: U59425)
97. SEQ ID NO: 96  
Proteinsequenz kodierend für für Glycine max 7S Samenglobulin
- 25 98. SEQ ID NO: 97  
Nukleinsäuresequenz kodierend für Zea mays 19kD Zein (GenBank Acc.-No.: E01144)
- 30 99. SEQ ID NO: 98  
Proteinsequenz kodierend für Zea mays 19kD Zein
99. SEQ ID NO: 99  
Nukleinsäuresequenz kodierend für Zea mays 19kD alpha Zein B1
- 35 100. (GenBank Acc.-No.: AF371269)
101. SEQ ID NO: 100  
Proteinsequenz kodierend für Zea mays 19kD alpha Zein B1
- 40 102. SEQ ID NO: 101  
Nukleinsäuresequenz kodierend für Zea mays 19kD alpha Zein B2 (GenBank Acc.-No.: AF371270)
103. SEQ ID NO: 102
- 45 104. Proteinsequenz kodierend für Zea mays 19kD alpha Zein B2

## 58

103. SEQ ID NO: 103  
Nukleinsäuresequenz kodierend für Zea mays 22kD alpha-zein  
(GenBank Acc.-No.: X61085)
- 5 104. SEQ ID NO: 104  
Proteinsequenz kodierend für Zea mays 22kD alpha-zein
105. SEQ ID NO: 105  
Nukleinsäuresequenz kodierend für Oryza sativa Prolamin  
10 (GenBank Acc.-No.: AB016503)
106. SEQ ID NO: 106  
Proteinsequenz kodierend für Oryza sativa Prolamin
- 15 107. SEQ ID NO: 107  
Nukleinsäuresequenz kodierend für A.sativa Avenin (GenBank  
Acc.-No.: M38446)
108. SEQ ID NO: 108  
20 Proteinsequenz kodierend für A.sativa Avenin
109. SEQ ID NO: 109  
Nukleinsäuresequenz kodierend für Hordeum vulgare C-Hordein  
(GenBank Acc.-No.: M36941)  
25
110. SEQ ID NO: 110  
Proteinsequenz Teil 1 kodierend für Hordeum vulgare C-Hordein
111. SEQ ID NO: 111  
30 Proteinsequenz Teil 2 kodierend für Hordeum vulgare C-Hordein
112. SEQ ID NO: 112  
Nukleinsäuresequenz kodierend für Triticum aestivum LMW Glu-  
tenin-1D1 (GenBank Acc.-No.: X13306)  
35
113. SEQ ID NO: 113  
Proteinsequenz kodierend für Triticum aestivum LMW Glute-  
nin-1D1
- 40 114. SEQ ID NO: 114  
Binärer Expressionsvektor für Agrobakterium vermittelte  
Pflanzentransformation pSUN2-USP.

115. SEQ ID NO: 115  
Partielle Nukleinsäuresequenz kodierend für Homogentisat-1,2-dioxygenase aus *Brassica napus* (HGD; EC-Nr.: 1.13.11.5)
- 5
116. SEQ ID NO: 116  
Nukleinsäuresequenz kodierend für Homogentisat-1,2-dioxygenase aus *Arabidopsis thaliana* (HGD; EC-Nr.: 1.13.11.5)
- 10 117. SEQ ID NO: 117  
Proteinsequenz kodierend für Homogentisat-1,2-dioxygenase aus *Arabidopsis thaliana* (HGD; EC-Nr.: 1.13.11.5)
118. SEQ ID NO: 118  
15 Nukleinsäuresequenz kodierend für Maleylacetoacetatisomerase aus *Arabidopsis thaliana* (MAAI; EC-Nr.: 5.2.1.2.)
119. SEQ ID NO: 119  
Proteinsequenz kodierend für Maleylacetoacetatisomerase aus  
20 *Arabidopsis thaliana* (MAAI; EC-Nr.: 5.2.1.2.)
120. SEQ ID NO: 120  
Nukleinsäuresequenz kodierend für Fumarylacetoacetathydrolase aus *Arabidopsis thaliana* (FAAH; EC-Nr.: 3.7.1.2)
- 25
121. SEQ ID NO: 121  
Proteinsequenz kodierend für Fumarylacetoacetathydrolase aus *Arabidopsis thaliana* (FAAH; EC-Nr.: 3.7.1.2)
- 30 122. SEQ ID NO: 122  
Nukleinsäuresequenz kodierend für Suppressionskonstrukt 2 (p3300.1-Toc159-GFP-RNAi)
- 35 123. SEQ ID NO: 123                      Oligonukleotidprimer OPN18
124. SEQ ID NO: 124                      Oligonukleotidprimer OPN19
125. SEQ ID NO: 125                      Oligonukleotidprimer OPN20
- 40 126. SEQ ID NO: 126                      Oligonukleotidprimer OPN21

#### Abbildungen

- 45 1. Fig.1: Schematische Darstellung der Speicherprotein-Suppressionskonstrukte.



## 60

Insert aus Vektor pCR2.1-sRNAi4 (1) (vgl. Beispiel 2d) und pCR2.1-sRNAi8 (2) (vgl. Beispiel 2e) kodierend für eine die AtCru3-, AtCRB und At2S3-Expression supprimierende dsRNA.

5 In den beiden Konstrukten sind die "sense"-Ribonukleotidsequenzen und die komplementären "antisense"-Ribonukleotidsequenzen (symbolisiert durch auf dem Kopf stehende Buchstaben) für die einzelnen zu supprimierenden Zielgene (AtCru3; AtCRB, At2S3) unterschiedlich angeordnet. Schraffierte Bereiche (I1,  
10 I2 etc.) stellen Intronsequenzen (Linker) dar.

2. Fig.2A-D: Symbolische Darstellung verschiedener dsRNAs in ihrer Sekundärstruktur.

15 S1, S2, ... S(n): "sense"-Ribonukleotidsequenz  
AS1, AS2, ... AS(n): "antisense"-Ribonukleotidsequenz  
I: Intronsequenz

Die Anordnung der einzelnen "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen kann so erfolgen, dass  
20 zunächst alle "sense"-Ribonukleotidsequenzen aneinander gereiht werden und so quasi einen "sense"-Strang bilden, worauf dann alle "antisense"-Ribonukleotidsequenzen aneinander zu einem "antisense"-Strang zusammengefügt werden (A und C).

25 Alternativ kann die Anordnung der einzelnen "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen so erfolgen, dass Paare von jeweils komplementären "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen an-  
30 einander gefügt werden (B und D).

"sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen können direkt aneinandergefügt sein (A und B) oder  
35 aber durch weitere Sequenzen beispielsweise Introns (I) voneinander getrennt sein (C und D).

3. Fig.3A-C: Symbolische Darstellung verschiedener dsRNAs in ihrer Sekundärstruktur.

40 S1, S2, ... S(n): "sense"-Ribonukleotidsequenz  
AS1, AS2, ... AS(n): "antisense"-Ribonukleotidsequenz  
SP: "SPACER"  
RE: Erkennungssequenz für Ribozym  
R: Ribozym  
45

## 61

"sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen können durch weitere Sequenzen ("SPACER"; SP) voneinander getrennt sein (A). Der Spacer kann beispielsweise eines Erkennungssequenz für ein Ribozym sein. Das korrespondierende Ribozym kann separat exprimiert werden (B) oder aber auch ebenfalls von dem Spacer kodiert sein (C).

4. Fig.4: Abbildung des Suppressionskonstrukts mit den entsprechenden Restriktionsenzymchnittstellen:

5. Fig.5A: Identifikation einer Pflanze, die den Albino-Phänotyp zeigt (links). Der Phänotyp ist identisch zur ppi2 Mutante, die Toc159 nicht mehr exprimieren kann. Als Kontrolle Pflanzen mit Wildtyp Phänotyp, die parallel gewachsen sind.

Fig. 5B: Fluoreszenz-Analyse der Pflanzen aus Fig.5A. Anregung der Fluoreszenz durch Licht im Wellenlängenbereich 470-490 nm. Es wurde dieselbe Vergrößerung gewählt wie in Fig.5A.

## Beispiele

## Allgemeine Methoden:

- 5 Alle Chemikalien, wenn nicht anders erwähnt, stammen von den Firmen Fluka (Buchs), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) und Sigma (Deisenhofen). Restriktionsenzyme, DNA-modifizierende Enzyme und Molekularbiologie-Kits wurden von den Firmen Amersham-Pharmacia (Freiburg), Biometra (Göttingen), Roche
- 10 (Mannheim), New England Biolabs (Schwalbach), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Qiagen (Hilden), Stratagen (Amsterdam, Niederlande), Invitrogen (Karlsruhe) und Ambion (Cambridgeshire, United Kingdom). Die verwendeten Reagenzien wurden entsprechend der Angaben des Herstellers eingesetzt.

- 15 Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die im Rahmen der vorliegenden Erfindung durchgeführten Klonierungsschritte wie z.B. Restriktionsspaltungen, Agarosegelelektrophorese, Reinigung von DNA-Fragmenten, Transfer von Nukleinsäuren auf Nitrozellulose und Nylonmembranen, Verknüpfen von DNA-Fragmenten, Transformation von E. coli Zellen, Anzucht von Bakterien, Vermehrung von Phagen und Sequenzanalyse rekombinanter DNA werden
- 20 wie bei Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6 beschrieben durchgeführt. Die Sequenzierung rekombinanter DNA-Moleküle erfolgt mit einem Laserfluoreszenz-DNA-Sequenzierer der Firma ABI nach der Methode von Sanger (Sanger et al. (1977) Proc Natl Acad Sci USA 74:5463-5467).
- 30

## Beispiel 1: Allgemeine Verfahren

- Die Pflanze *Arabidopsis thaliana* repräsentiert ein Mitglied der höheren Pflanzen (Samenpflanzen). Diese Pflanze ist eng verwandt mit anderen Pflanzenarten aus der Familie der Cruciferen wie z.B. *Brassica napus*, aber auch mit anderen Pflanzenfamilien der Dikotyledonen. Aufgrund des hohen Grades an Homologie ihrer DNA-Sequenzen bzw. Polypeptidsequenzen kann *Arabidopsis thaliana* als
- 40 Modellpflanze für andere Pflanzenarten eingesetzt werden.

a) Anzucht von *Arabidopsis* Pflanzen

- Die Pflanzen werden entweder auf Murashige-Skoog Medium mit 0,5 % Saccharose (Ogas et al. (1997) Science 277:91-94) oder auf Erde gezogen (Focks & Benning (1998) Plant Physiol 118:91-101). Um einheitliche Keimungs- und Blühzeiten zu erreichen, werden die

## 63

Samen nach Ausplattieren bzw. Ausstreuen auf Erde zwei Tage bei 4° C stratifiziert. Nach der Blüte werden die Schoten markiert. Entsprechend der Markierungen werden dann Schoten mit einem Alter von 6 bis 20 Tagen nach der Blüte geerntet.

5

b) Isolierung von total RNA und poly-A<sup>+</sup> RNA aus Pflanzen

Für die Herstellung von Suppressionskonstrukten wird RNA bzw. polyA<sup>+</sup> RNA isoliert. RNA wurde aus Schoten von Arabidopsis Pflanzen nach folgender Vorschrift isoliert: Schotenmaterial im Alter von 6 bis 20 Tage nach Blüte wurde geerntet und in flüssigem Stickstoff schockgefroren. Das Material wurde vor der weiteren Verwendung bei -80°C gelagert. 75 mg des Materials wurde im gekühlten Mörser zu einem feinem Pulver gemahlen und mit 200 µL des Lysis-Puffers aus dem Ambion RNAqueos-Kit versetzt. Die Isolierung der totalen RNA wurde dann nach Herstellerangaben durchgeführt. Die RNA wurde mit 50 µL Elutionspuffer (Ambion) eluiert und die Konzentration durch Absorption einer 1 zu 100 verdünnten Lösung am Photometer (Eppendorf) bei 260 nm bestimmt. 40 µg/ml RNA entspricht dabei einer Absorption von 1. Die RNA-Lösungen wurden mit RNase freiem Wasser auf eine Konzentration von 1 µg/µL eingestellt. Die Konzentrationen wurden durch Agarosegelelektrophorese überprüft. Zur Isolierung von polyA<sup>+</sup> RNA wurde oligo(dT)-Zellulose von Amersham Pharmacia nach Herstellerangaben verwendet. RNA bzw. polyA<sup>+</sup> RNA wurde bei -70°C gelagert.

## c) Konstruktion der cDNA-Bank

Zur Konstruktion der cDNA-Bank aus Arabidopsis Schoten-RNA wurde die Erststrangsynthese unter Verwendung von Reverser Transkriptase aus Maus-Leukämie-Virus (Clontech) und Oligo-d(T)-Primern, die Zweitstrangsynthese durch Inkubation mit DNA-Polymerase I, Klenow-Enzym und RNase H-Spaltung bei 12°C (2 Std.), 16°C (1 Std.) und 22°C (1 Std.) erzielt. Die Reaktion wurde durch Inkubation bei 65°C (10 min) gestoppt und anschließend auf Eis überführt. Doppelsträngige DNA-Moleküle wurde mit T4-DNA-Polymerase (Roche, Mannheim) bei 37°C (30 min) mit glatten Enden versehen. Die Nukleotide wurden durch Phenol/Chloroform-Extraktion und Sephadex-G50-Zentrifugiersäulen entfernt. EcoRI/XhoI-Adapter (Pharmacia, Freiburg, Deutschland) wurden mittels T4-DNA-Ligase (Roche, 12°C, über Nacht) an die cDNA-Enden ligiert, mit XhoI nachgeschnitten und durch Inkubation mit Polynukleotidkinase (Roche, 37°C, 30 min) phosphoryliert. Dieses Gemisch wurde der Trennung auf einem Low-Melting-Agarose-Gel unterworfen. DNA-Moleküle über 300 Basenpaaren wurden aus dem Gel eluiert, Phenol-extrahiert, auf Elutip-D-Säulen (Schleicher und Schüll, Dassel, Deutschland) konzentriert und an Vektorarme ligiert und in lambda-ZAPII-Phagen

oder lambda-ZAP-Express-Phagen unter Verwendung des Gigapack Gold-Kits (Stratagene, Amsterdam, Niederlande) verpackt, wobei Material des Herstellers verwendet und seine Anweisungen befolgt wurden.

5

- d) Isolierung von genomischer DNA aus Pflanzen wie *Arabidopsis thaliana* oder *Brassica napus* (CTAB-Methode)

Zur Isolierung genomischer DNA aus Pflanzen wie *Arabidopsis thaliana* oder *Brassica napus* werden ca. 0,25 g Blattmaterial junger Pflanzen im vegetativen Stadium in flüssigem Stickstoff zu feinem Pulver gemörsert. Das pulverisierte Pflanzenmaterial wird zusammen mit 1 ml 65°C-warmem CTAB I-Puffer (CTAB: Hexadecyltrimethylammoniumbromid, auch genannt Cetyltrimethylammoniumbromid; Sigma Kat.-Nr.: H6269) und 20 µl β-Mercaptoethanol in einen vorgewärmten zweiten Mörser gegeben und nach vollständiger Homogenisierung wird der Extrakt in ein 2 ml Eppendorf-Gefäß überführt und für 1 h bei 65°C unter regelmäßiger, vorsichtiger Durchmischung inkubiert. Nach Abkühlung auf Raumtemperatur wird der Ansatz mit 1 ml Chloroform/Octanol (24:1, mit 1M Tris/HCl, pH 8,0 ausgeschüttelt) durch langsames Invertieren extrahiert und zur Phasentrennung für 5 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Anschließend wird die wässrige Phase erneut mit 1 ml Chloroform/Octanol extrahiert, zentrifugiert und durch Invertieren mit 1/10 Volumen auf 65°C vorgewärmtem CTAB II-Puffer sorgfältig gemischt. Anschließend wird der Ansatz durch vorsichtiges Schwenken mit 1 ml Chloroform/Octanol-Gemisch (siehe oben) versetzt und zur erneuten Phasentrennung für 5 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Die wässrige untere Phase wird in ein frisches Eppendorf-Gefäß überführt und die obere organische Phase wird in einem frischen Eppendorf-Gefäß erneut für 15 min bei 8,500 rpm (7,500 x g) und Raumtemperatur zentrifugiert. Die hieraus resultierende wässrige Phase wird mit der wässrigen Phase des vorherigen Zentrifugationsschrittes vereinigt und der gesamte Ansatz mit exakt demselben Volumen vorgewärmtem CTAB III-Puffer versetzt. Es folgt eine Inkubation bei 65°C, bis die DNA in Flocken ausfällt. Dies kann bis zu 1 h dauern oder durch Inkubation bei 37°C über Nacht erfolgen. Das aus dem anschließenden Zentrifugationsschritt (5 min, 2000 rpm (500 x g), 4°C) resultierende Sediment wird mit 250 µl auf 65°C vorgewärmtem CTAB IV-Puffer versetzt und für mindestens 30 min bzw. bis zur vollständigen Auflösung des Sediments bei 65°C inkubiert. Anschließend wird die Lösung zur Fällung der DNA mit 2,5 Volumina eiskaltem Ethanol vermischt und für 1h bei -20°C inkubiert. Alternativ kann der Ansatz mit 0.6 Volumina Isopropanol vermischt und ohne weitere Inkubation sofort für 15 min bei 8,500 rpm (7,500 x g) und 4°C zentrifugiert werden. Die sedimentierte DNA wird durch

## 65

Invertieren des Eppendorf-Gefäßes zweimal mit je 1 ml 80%igem eiskaltem Ethanol gewaschen, nach jedem Waschschrift erneut zentrifugiert (5 min, 8,500 rpm (7,500 x g), 4°C) und anschließend für ca. 15 min luftgetrocknet. Abschließend wird die DNA in

- 5 100 µl TE mit 100 µg/ml RNase resuspendiert und für 30 min bei Raumtemperatur inkubiert. Die DNA Lösung ist nach einer weiteren Inkubationsphase über Nacht bei 4°C homogen und kann für weiterführende Experimente verwendet werden.

## 10 Lösungen für CTAB:

Lösung I (für 200 ml):

- 100 mM Tris/HCl pH 8,0 (2,42 g)  
1,4 M NaCl (16,36 g)  
15 20 mM EDTA (8,0 ml von 0,5 M Stammlösung)  
2 % (w/v) CTAB (4,0 g)

Jeweils vor der Verwendung werden frisch zugesetzt:  
2 % β-Mercaptoethanol (20 µl für 1 ml Lösung I).

20

Lösung II (für 200 ml):

- 0,7 M NaCl (8,18 g)  
10 % (w/v) CTAB (20 g)

## 25 Lösung III (für 200 ml):

- 50 mM Tris/HCl pH 8,0 (1,21 g)  
10 mM EDTA (4 ml 0,5 M von 0,5 M Stammlösung)  
1 % (w/v) CTAB (2,0 g)

## 30 Lösung IV (High-salt TE) (für 200 ml):

- 10 mM Tris/ HCl pH 8,0 (0,242 g)  
0,1 mM EDTA (40 µl 0,5 M Stammlösung)  
1 M NaCl (11, 69 g)

## 35 Chloroform/Octanol (24:1) (für 200 ml):

- 192 ml Chloroform  
8 ml Octanol

Die Mischung wird 2x mit 1 M TrisHCl pH 8,0 ausgeschüttelt und vor Licht geschützt gelagert.

40

## Beispiel 2: Herstellung von Suppressionskonstrukten

Ausgehend von der genomischer Arabidopsis thaliana DNA oder cDNA wurden über PCR mittels der aufgeführten Oligonukleotide folgende

- 45 Fragmente von Speicherproteinsequenzen amplifiziert. Dabei kam nachfolgendes PCR Protokoll zum Einsatz:

## 66

Zusammensetzung des PCR-Ansatzes (50 µL):

- 5,00 µL Template cDNA oder genomische DNA (ca. 1 µg)
- 5,00 µL 10x Puffer (Advantage-Polymerase) + 25 mM MgCl<sub>2</sub>
- 5 5,00 µL 2mM dNTP
- 1,25 µL je Primer (10 pmol/µL)
- 0,50 µL Advantage-Polymerase (Clontech)

PCR-Programm: Anfangsdenaturierung für 2 min bei 95°C, dann 35 Zyklen mit 45 sec 95°C, 45 sec 55°C und 2 min 72°C. Abschliessende Extension von 5 min bei 72°C.

a) Ausgangsvektor pCR2.1-AtCRU3-RNAi

- 15 Aus genomischer *Arabidopsis thaliana* DNA wird mit nachfolgendem Oligonukleotid-Primerpaar ein Exonbereich mit dem vollständigen anschließenden Intron einschliesslich der an das Intron anschließenden Spleiß-Akzeptorsequenz des 12S Speicherprotein AtCRU3 (Basenpaar 1947 bis 2603 der Sequenz mit der GenBank Acc.-No: 20 U66916) amplifiziert:

ONP1 (SEQ ID NO: 134):

5'-ATAAGAATGCGGCCGCGTGTTCATTTGGCCGGAACAAC-3'

- 25 ONP2 (SEQ ID NO: 135):

5'-CCCGGATCCTTCTGTAAACATTTGACAAAACATG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-1 30 Vektor und die Sequenz überprüft.

Für die den antisense-Strang der dsRNA kodierende Sequenz wird aus *Arabidopsis thaliana* cDNA lediglich das gleiche Exon wie oben (Basenpaar 1947 bis 2384) mit dem nachfolgenden Primerpaar amplifiziert: 35

ONP3 (SEQ ID NO: 136):

5'ATAAGAATGCGGCCGCGTGTTCATTTGGCCGGAACAAC-3'

- 40 ONP4 (SEQ ID NO: 137):

5' ATAAGAATGCGGCCGCGGATCCACCCTGGAGAACGCCACGAGTG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-2 45 Vektor und die Sequenz überprüft.

## 67

- 0,5 µg von Vektor pCR2.1-1 werden mit dem Restriktionsenzym BamHI (New England Biolabs) für 2 Stunden nach Herstellerangaben inkubiert und dann für 15 min mit alkalischer Phosphatase (New England Biolabs) dephosphoryliert. Der so präparierte Vektor (1 µL) wird dann mit dem aus Vektor pCR2.1-2 gewonnenen Fragment ligiert. Dazu werden 0,5 µg von Vektor pCR2.1-2 2 Stunden mit BamHI (New England Biolabs) verdaut und die DNA-Fragmente per Gelelektrophorese aufgetrennt. Das neben dem Vektor (3,9 kb) entstandene 489 bp grosse Stück wird aus dem Gel ausgeschnitten und mit dem "Gelpurification"-Kit (Qiagen) nach Herstellerangaben aufgereinigt und mit 50 µL Elutionspuffer eluiert. 10 µL des Eluates werden mit Vektor pCR2.1-1 (s.o.) über Nacht bei 16°C ligiert (T4 Ligase, New England Biolabs). Die Ligationsprodukte werden dann in TOP10 Zellen (Stratagene) nach Herstellerangaben transformiert und entsprechend selektioniert. Positive Klone werden mit dem Primerpaar ONP1 und ONP2 durch PCR verifiziert. Der erhaltene Vektor wird pCR2.1-AtCRU3-RNAi genannt. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 105 beschrieben.

20 b) Ausgangsvektor pCR2.1-AtCRB-RNAi

Mit nachfolgendem Oligonukleotid-Primerpaar wird ein Exonbereich des 12S Speicherprotein AtCRB (SEQ ID NO: 117 bzw. 118; Basenpaar 601 bis 1874 der Sequenz mit der GenBank Acc.-No: M37248) aus

25 Arabidopsis thaliana cDNA amplifiziert:

ONP5 (SEQ ID NO: 138):

5' ATAAGAATGCGGCCGCGGATCCCTCAGGGTCTTTTCTTGCCCACT-3'

30 ONP6 (SEQ ID NO: 139):

5'-CCGCTCGAGTTTACGGATGGAGCCACGAAG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-3

35 Vektor und die Sequenz überprüft.

- Für den als Linker fungierenden Bereich wird aus Arabidopsis thaliana genomischer DNA ein Intron mit den entsprechenden Spliceakzeptor und -donorsequenzen der flankierenden Exons (Basenpaar 1874 bis 2117 der Sequenz mit der GenBank Acc.-No: M37248) mit dem nachfolgenden Primerpaar amplifiziert:

ONP7 (SEQ ID NO: 140):

5'-CCGCTCGAGGTAAGCTCAACAAATCTTTAG-3'

45 ONP8 (SEQ ID NO: 141):

5'-ACGCGTCGACGCGTTCTGCGTGCAAGATATT-3'



## 68

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-4 Vektor und die Sequenz überprüft.

- 5 Das Konstrukt für AtCRB wird in einer ähnlichen Strategie wie für AtCRU3 erläutert, erstellt. Vektor pCR2.1-3 wird mit mit XhoI (New England Biolabs) für 2 Stunden inkubiert und dephosphoryliert (alkalische Phosphatase, New England Biolabs). Vektor pCR2.1-4 wird ebenfalls mit XhoI in derselben Weise inkubiert und  
 10 die Gelfragmente per Gelelektrophorese aufgetrennt. Die entsprechenden Fragmente werden in der unter AtCRU3 beschriebenen Art und Weise aufgereinigt und ligiert, resultierend nach Bakterien-transformation in dem Vektor pCR2.1-AtCRB Exon/Intron. Dieser Vektor wird für 2 Stunden mit XbaI (NEB), anschliessend für 15  
 15 min mit Klenow-Fragment (NEB), dann für 2 Stunden mit SalI inkubiert und zuletzt 15 min mit alkalischer Phosphatase (NEB) behandelt. Parallel wird der Vektor pCR2.1-3 mit BamHI (NEB), dann 15 min mit Klenow-Fragment und anschliessend 2 Stunden mit XhoI (NEB) inkubiert. Das Exon-Fragment von AtCRB wird nach Gelelektrophorese isoliert, gereinigt und zur Ligation eingesetzt. Beide  
 20 Fragmente wurden dann ligiert und der Vektor pCR2.1-AtCRB-RNAi resultierte. Der erhaltene Vektor wird pCR2.1-AtCRB-RNAi genannt. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 107 beschrieben.

- 25 c) Ausgangsvektor pCR2.1-At2S3-RNAi

Mit nachfolgendem Oligonukleotid-Primerpaar wird ein Exonbereich des 2S Speicherprotein At2S3 (SEQ ID NO: 3 bzw. 4; Basenpaar 212 bis 706 der Sequenz mit der GenBank Acc.-No: M22035) amplifi-  
 30 ziert:

ONP9 (SEQ ID NO: 142):

5'-ATAAGAATGCGGCCGCGGATCCATGGCTAACAAGCTCTTCCTCGTC-3'

- 35 ONP10 (SEQ ID NO: 143):

5'-ATAAGAATGCGGCCGCGGATCCCTAGTAGTAAGGAGGGAAGAAAG-3'

Das PCR-Produkt wird in den pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in dem pCR2.1-5  
 40 Vektor und die Sequenz überprüft. Für den als Linker fungierenden Bereich wird das gleiche Intron wie unter b) mit den Primern OPN 7 und OPN 8 amplifiziert eingesetzt.

45 Das Konstrukt für At2S3 wird in einer ähnlichen Strategie wie für AtCRU3 erläutert, erstellt. Vektor pCR2.1-5 wird mit mit XhoI (New England Biolabs) für 2 Stunden inkubiert und dephosphoryliert (alkalische Phosphatase, New England Biolabs). Vektor

## 69

pCR2.1-3 werden ebenfalls mit XhoI in derselben Weise inkubiert und die Gelfragmente per Gelelektrophorese aufgetrennt. Die entsprechenden Fragmente werden in der unter AtCRU3 beschriebenen Art und Weise aufgereinigt und ligiert, resultierend nach Bakterientransformation in dem Vektor pCR2.1-At2S3 Exon/Intron. Dieser Vektor wird für 2 Stunden mit SalI (NEB), anschliessend für 15 min mit Klenow-Fragment (NEB) inkubiert und zuletzt 15 min mit alkalischer Phosphatase (NEB) behandelt. Parallel wird der Vektor pCR2.1-5 mit BamHI (NEB) und dann 15 min mit Klenow-Fragment inkubiert. Das Exon-Fragment von At2S3 wird nach Gelelektrophorese isoliert, gereinigt und zur Ligation eingesetzt. Beide Fragmente werden dann ligiert und der Vektor pCR2.1-At2S3-RNAi resultiert. Die für die dsRNA kodierende Nukleinsäuresequenz ist durch SEQ ID NO: 109 beschrieben.

15

## d) Herstellung von Super-Suppressionskonstrukt 1

Die Vektoren pCR2.1-AtCRU3-RNAi und pCR2.1-4 (siehe oben) werden mit den Restriktionsenzymen XhoI und SalI für 2 Stunden bei 37°C inkubiert, die DNA-Fragmente durch Agarose-Gelelektrophorese aufgetrennt und sowohl der Vektor als auch das PCR-Insert aus pCR2.1-4 ausgeschnitten und mit dem "Gelpurification"-Kit von Qiagen nach Herstellerangaben aufgereinigt und mit 50 µL Elutionspuffer eluiert. Vom Vektor wird 1 µL, vom PCR-Insert aus pCR2.1-4 8 µL der Eluate für die Ligation eingesetzt, resultierend in dem Konstrukt pCR2.1-sRNAi1. Dieser Vektor wird für 2 Stunden mit dem Restriktionsenzym XhoI und dann für 15 min mit Klenow-Fragment inkubiert.

Der Vektor pCR2.1-AtCRB-RNAi (siehe oben) wird mit dem Enzym EcoRI für 2 Stunden inkubiert und ebenfalls 15 min mit Klenow-Fragment behandelt. Beide Inkubationsansätze werden durch Gelelektrophorese aufgetrennt und jeweils der Vektor (pCR2.1-sRNAi1) bzw. das Insert (aus pCR2.1-AtCRB-RNAi) aus dem Agarosegel ausgeschnitten und die DNA-Fragmente wie oben beschrieben aufgereinigt. Für die Ligation werden 1 µL des Eluates vom Vektor und 8 µL des Eluates vom Insert eingesetzt und bei 4°C über Nacht inkubiert. Das resultierende Konstrukt wird mit pCR2.1-sRNAi2 bezeichnet. Der resultierende Vektor wird mit dem Enzym XbaI und anschliessend mit Klenow-Fragment inkubiert. Der Vektor pCR2.1-4 wird mit den Enzymen EcoRV und XbaI und anschliessend mit Klenow-Fragment inkubiert. Nach Gelelektrophorese und -reinigung wird das Fragment aus pCR2.1-4 mit dem Vektor pCR2.1-sRNAi2 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi3. Der resultierende Vektor wird dann mit dem Enzym ApaI für 2 Stunden und dann mit Klenow-Fragment für 15 min inkubiert. Als Insert wird der Vektor pCR2.1-At2S3-RNAi mit dem Enzym EcoRI für 2 Stunden und dann mit

## 70

Klenow-Fragment für 15 min inkubiert. Nach Gelelektrophorese und -reinigung werden die Eluate ligiert, resultierend in dem Vektor pCR2.1-sRNAi4. Aus diesem Vektor wird dann das sRNAi4-Fragment (SEQ ID NO: 144; vgl. Fig. 1(1)), kodierend für die super-supprimierende dsRNA, durch Inkubation mit HindIII und PvuI ausgeschnitten und in den binären Vektor pSUN-USP (SEQ ID NO: 179) ligiert. Das Konstrukt dient der gleichzeitigen Suppression von Arabidopsis thaliana Speicherproteinen CRB (SEQ ID NO: 4), CRU3 (SEQ ID NO: 112) und At2S3 (SEQ ID NO: 118).

10

Der verwendete Vektor pSUN-USP ist ein binärer Vektor zur Pflanzentransformation auf Basis von pBinAR (Höfgen und Willmitzer (1990) Plant Science 66: 221-230). Eine gewebespezifische Expression im Samen läßt sich unter Verwendung des gewebespezifischen Promotors USP-Promotors erzielt.

e) Herstellung von Super-Suppressionskonstrukt 2

Ausgehend von Arabidopsis thaliana cDNA wird ein Fragment aus dem Speicherprotein AtCRU3 (SEQ ID NO: 111, 112) mit dem nachfolgenden Oligonukleotid-Primerpaar unter den in Beispiel 2 angegebenen PCR-Bedingungen amplifiziert:

OPN 11: 5'-AAAAGGCCTGTGTTCCATTTGGCCGGAACAAC-3' (SEQ ID NO: 148)

25 OPN 12: 5'-AAAGATATCACCTGGAGAACGCCACGAGTG-3' (SEQ ID NO: 149).

Das erhaltene Fragment wird in den Vektor pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in den pCR2.1-6 und die Sequenzen überprüft.

30

Ausgehend von Arabidopsis thaliana cDNA wird ein Fragment aus dem Speicherprotein At2S3 (SEQ ID NO: 3, 4) mit dem nachfolgenden Oligonukleotid-Primerpaar unter den in Beispiel 2 angegebenen PCR-Bedingungen amplifiziert:

35

OPN 13: 5'-AAAAGGCCTATGGCTAACAAGCTCTTCCTCGTC-3' (SEQ ID NO: 150)

OPN 14: 5'-AAAGATATCCTAGTAGTAAGGAGGGAAGAAAG-3' (SEQ ID NO: 151).

Das erhaltene Fragment wird in den Vektor pCR2.1-TOPO Vektor (Invitrogen) gemäss Herstellerangaben kloniert, resultierend in den pCR2.1-7 und die Sequenzen überprüft.

40

Aus den pCR2.1-3, pCR2.1-4 (siehe Beispiel 2) und pCR2.1-6 und pCR2.1-7 werden dann die Konstrukte folgendermassen zusammen ligiert: Der Vektor pCR2.1-3 wird 2 Stunden mit EcoRV inkubiert und anschliessend 15 min mit alkalischer Phosphatase dephosphoryliert. Der Vektor pCR2.1-6 wird mit den Enzymen StuI und EcoRV

45

## 71

für 2 Stunden inkubiert und das PCR-Insert über Gelelektrophorese und -reinigung isoliert. Vektor pCR2.1-3 und Insert aus pCR2.1-6 werden dann über Nacht bei 4°C ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi5. Dieser Vektor wird dann mit EcoRV inkubiert und dephosphoryliert und mit dem StuI/ EcoRV inkubierten und ge-  
 5 laufigereinigten Fragment aus pCR2.1-7 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi6. Dieser Vektor wird dann mit XhoI inkubiert und dephosphoryliert. Der Vektor pCR2.1-4 wird mit SalI und XhoI inkubiert und das Insert aus pCR2.1-4 mit dem vorberei-  
 10 teten Vektor pCR2.1-sRNAi6 ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi7. Ausgehend von pCR2.1-sRNAi7 wird eine PCR mit den nachfolgenden Primerpaar unter den in Beispiel 2 gegebenen Bedingungen durchgeführt:

15 OPN 15: 5' CCGCTCGAGCTCAGGGTCTTTTCTTGCCCACT (SEQ ID NO: 152)  
 OPN 16: 5'-CCGGTCGACCTAGTAGTAAGGAGGAAGAAAG (SEQ ID NO: 153).

Das resultierende PCR-Produkt wird mit den Enzymen XhoI und SalI inkubiert. Das Fragment wird dann in den Vektor pCR2.1-sRNAi7  
 20 (inkubiert mit XhoI) ligiert, resultierend in dem Konstrukt pCR2.1-sRNAi8. Aus diesem Vektor wird dann das sRNAi8-Fragment (SEQ ID NO: 146; vgl. Fig. 1(2)), kodierend für die super-supprimierende dsRNA, durch Inkubation mit HindIII und XbaI ausgeschnitten und in den binären Vektor pSUN-USP (SEQ ID NO: 179) li-  
 25 giert. Das Konstrukt dient der gleichzeitigen Suppression von Arabidopsis thaliana Speicherproteinen CRB (SEQ ID NO: 4), CRU3 (SEQ ID NO: 112) und At2S3 (SEQ ID NO: 118).

## Beispiel 3: Transformation von Agrobacterium

30 Die Agrobacterium-vermittelte Pflanzentransformation kann zum Beispiel unter Verwendung der Agrobacterium tumefaciens-Stämme GV3101 (pMP90) (Koncz und Schell (1986) Mol Gen Genet 204: 383-396) oder LBA4404 (Clontech) durchgeführt werden. Die Transfor-  
 35 mation kann durch Standard-Transformationstechniken durchgeführt werden (Deblaere et al. (1984) Nucl Acids Res 13:4777-4788).

## Beispiel 4: Pflanzentransformation

40 Die Agrobacterium-vermittelte Pflanzentransformation kann unter Verwendung von Standard-Transformations- und Regenerationstechniken durchgeführt werden (Gelvin, Stanton B., Schilperoort, Robert A., Plant Molecular Biology Manual, 2. Aufl., Dordrecht: Kluwer Academic Publ., 1995, in Sect., Ringbuc Zentrale Signatur: BT11-P  
 45 ISBN 0-7923-2731-4; Glick, Bernard R., Thompson, John E., Methods

in Plant Molecular Biology and Biotechnology, Boca Raton: CRC Press, 1993, 360 S., ISBN 0-8493-5164-2).

- Die Transformation mittels Agrobacterium von Arabisopsis thaliana  
5 wird durch die Methode nach Bechthold et al., 1993 (C.R. Acad. Sci. Ser. III Sci. Vie., 316, 1194-1199) durchgeführt.  
Beispielsweise kann Raps mittels Kotyledonen- oder Hypokotyl-  
transformation transformiert werden (Moloney et al., Plant Cell  
Report 8 (1989) 238-242; De Block et al., Plant Physiol. 91  
10 (1989) 694-701). Die Verwendung von Antibiotika für die Agrobac-  
terium- und Pflanzenselektion hängt von dem für die Transforma-  
tion verwendeten binären Vektor und Agrobacterium-Stamm ab. Die  
Rapsselektion wird gewöhnlich unter Verwendung von Kanamycin als  
selektierbarem Pflanzenmarker durchgeführt.
- 15 Der Agrobacterium-vermittelte Gentransfer in Lein (Linum usita-  
tissimum) läßt sich unter Verwendung von beispielsweise einer von  
Mlynarova et al. (1994) Plant Cell Report 13:282-285 beschriebe-  
nen Technik durchführen.
- 20 Die Transformation von Soja kann unter Verwendung von beispiels-  
weise einer in EP-A-0 0424 047 (Pioneer Hi-Bred International)  
oder in EP-A-0 0397 687, US 5,376,543, US 5,169,770 (University  
Toledo) beschriebenen Technik durchgeführt werden.
- 25 Die Pflanzentransformation unter Verwendung von Teilchenbe-  
schuß, Polyethylenglycol-vermittelter DNA-Aufnahme oder über die  
Siliziumcarbonatfaser-Technik ist beispielsweise beschrieben  
von Freeling und Walbot "The maize handbook" (1993) ISBN  
30 3-540-97826-7, Springer Verlag New York).

Beispiel 5: Untersuchung der Expression eines rekombinanten  
Genproduktes in einem transformierten Organismus

- 35 Die Aktivität eines rekombinanten Genproduktes im transformierten  
Wirtsorganismus wurde auf der Transkriptions- und/oder der  
Translationsebene gemessen.

- Ein geeignetes Verfahren zur Bestimmung der Menge an Transkrip-  
40 tion des Gens (ein Hinweis auf die Menge an RNA, die für die  
Tranlation des Genproduktes zur Verfügung steht) ist die Durch-  
führung eines Northern-Blots wie unten ausgeführt (als Bezugs-  
stelle siehe Ausubel et al. (1988) Current Protocols in Molecular  
Biology, Wiley: New York, oder den oben erwähnten Beispielteil),  
45 wobei ein Primer, der so gestaltet ist, daß er an das Gen von In-  
teresse bindet, mit einer nachweisbaren Markierung (gewöhnlich  
radioaktiv oder chemilumineszent) markiert wird, so daß, wenn die

## 73

Gesamt-RNA einer Kultur des Organismus extrahiert, auf einem Gel aufgetrennt, auf eine stabile Matrix transferiert und mit dieser Sonde inkubiert wird, die Bindung und das Ausmaß der Bindung der Sonde das Vorliegen und auch die Menge der mRNA für dieses Gen anzeigt. Diese Information zeigt den Grad der Transkription des transformierten Gens an. Zelluläre Gesamt-RNA kann aus Zellen, Geweben oder Organen mit mehreren Verfahren, die alle im Fachgebiet bekannt sind, wie zum Beispiel das von Bormann, E.R., et al. (1992) Mol. Microbiol. 6:317-326 beschriebene, präpariert werden.

10

## Northern-Hybridisierung:

Für die RNA-Hybridisierung wurden 20 µg Gesamt-RNA oder 1 µg poly(A)<sup>+</sup>-RNA mittels Gelelektrophorese in Agarosegelen mit einer Stärke von 1,25 % unter Verwendung von Formaldehyd, wie beschrieben in Amasino (1986, Anal. Biochem. 152, 304) aufgetrennt, mittels Kapillaranziehung unter Verwendung von 10 x SSC auf positiv geladene Nylonmembranen (Hybond N+, Amersham, Braunschweig) übertragen, mittels UV-Licht immobilisiert und 3 Stunden bei 68°C unter Verwendung von Hybridisierungspuffer (10 % Dextransulfat Gew./Vol., 1 M NaCl, 1 % SDS, 100 mg Heringssperma-DNA) vorhybridisiert. Die Markierung der DNA-Sonde mit dem Highprime DNA labeling-Kit (Roche, Mannheim, Deutschland) erfolgte während der Vorhybridisierung unter Verwendung von alpha-<sup>32</sup>P-dCTP (Amersham Pharmacia, Braunschweig, Deutschland). Die Hybridisierung wurde nach Zugabe der markierten DNA-Sonde im gleichen Puffer bei 68°C über Nacht durchgeführt. Die Waschschrte wurden zweimal für 15 min unter Verwendung von 2 X SSC und zweimal für 30 min unter Verwendung von 1 X SSC, 1 % SDS, bei 68°C durchgeführt. Die Exposition der verschlossenen Filter wurde bei -70°C für einen Zeitraum von 1 bis 14 T durchgeführt.

Zur Untersuchung des Vorliegens oder der relativen Menge an von dieser mRNA translatiertem Protein können Standardtechniken, wie ein Western-Blot, eingesetzt werden (siehe beispielsweise Ausubel et al. (1988) Current Protocols in Molecular Biology, Wiley: New York). Bei diesem Verfahren werden die zellulären Gesamt-Proteine extrahiert, mittels Gelelektrophorese aufgetrennt, auf eine Matrix, wie Nitrozellulose, übertragen und mit einer Sonde, wie einem Antikörper, der spezifisch an das gewünschte Protein bindet, inkubiert. Diese Sonde ist gewöhnlich mit einer chemilumineszenten oder kolorimetrischen Markierung versehen, die sich leicht nachweisen läßt. Das Vorliegen und die Menge der beobachteten Markierung zeigt das Vorliegen und die Menge des gewünschten, in der Zelle vorliegenden mutierten Proteins an.

45

## 74

Beispiel 6: Analyse der Auswirkung der rekombinanten Proteine auf die Produktion des gewünschten Produktes

Die Auswirkung der genetischen Modifikation in Pflanzen, Pilzen,  
5 Algen, Ciliaten oder auf die Produktion einer gewünschten Verbindung (wie einer Fettsäure) kann bestimmt werden, indem die modifizierte Mikroorganismen oder die modifizierte Pflanze unter geeigneten Bedingungen (wie den vorstehend beschriebenen) gezüchtet werden und das Medium und/oder die zellulären Komponenten auf die  
10 erhöhte Produktion des gewünschten Produktes (d.h. von Lipiden oder einer Fettsäure) untersucht wird. Diese Analysetechniken sind dem Fachmann bekannt und umfassen Spektroskopie, Dünnschichtchromatographie, Färbeverfahren verschiedener Art, enzymatische und mikrobiologische Verfahren sowie analytische Chromatographie, wie Hochleistungs-Flüssigkeitschromatographie (siehe  
15 beispielsweise Ullman, Encyclopedia of Industrial Chemistry, Bd. A2, S. 89-90 und S. 443-613, VCH: Weinheim (1985); Fallon, A., et al., (1987) "Applications of HPLC in Biochemistry" in: Laboratory Techniques in Biochemistry and Molecular Biology, Bd. 17; Rehm et al. (1993) Biotechnology, Bd. 3, Kapitel III: "Product recovery and purification", S. 469-714, VCH: Weinheim; Belter, P.A., et al. (1988) Bioseparations: downstream processing for Biotechnology, John Wiley and Sons; Kennedy, J.F., und Cabral, J.M.S. (1992) Recovery processes for biological Materials, John Wiley  
20 and Sons; Shaeiwitz, J.A., und Henry, J.D. (1988) Biochemical Separations, in: Ullmann's Encyclopedia of Industrial Chemistry, Bd. B3; Kapitel 11, S. 1-27, VCH: Weinheim; und Dechow, F.J. (1989) Separation and purification techniques in biotechnology, Noyes Publications).

30 Neben den oben erwähnten Verfahren werden Pflanzenlipide aus Pflanzenmaterial wie von Cahoon et al. (1999) Proc. Natl. Acad. Sci. USA 96 (22):12935-12940, und Browse et al. (1986) Analytic Biochemistry 152:141-145, beschrieben extrahiert. Die qualitative  
35 und quantitative Lipid- oder Fettsäureanalyse ist beschrieben bei Christie, William W., Advances in Lipid Methodology, Ayr/Scotland: Oily Press (Oily Press Lipid Library; 2); Christie, William W., Gas Chromatography and Lipids. A Practical Guide - Ayr, Scotland: Oily Press, 1989, Repr. 1992, IX, 307 S. (Oily  
40 Press Lipid Library; 1); "Progress in Lipid Research, Oxford: Pergamon Press, 1 (1952) - 16 (1977) u.d.T.: Progress in the Chemistry of Fats and Other Lipids CODEN.

Zusätzlich zur Messung des Endproduktes der Fermentation ist  
45 es auch möglich, andere Komponenten der Stoffwechselwege zu analysieren, die zur Produktion der gewünschten Verbindung verwendet werden, wie Zwischen- und Nebenprodukte, um die Gesamteffizienz

## 75

der Produktion der Verbindung zu bestimmen. Die Analyseverfahren umfassen Messungen der Nährstoffmengen im Medium (z.B. Zucker, Kohlenwasserstoffe, Stickstoffquellen, Phosphat und andere Ionen), Messungen der Biomassezusammensetzung und des Wachstums, 5 Analyse der Produktion üblicher Metabolite von Biosynthesewegen und Messungen von Gasen, die während der Fermentation erzeugt werden. Standardverfahren für diese Messungen sind in Applied Microbial Physiology; A Practical Approach, P.M. Rhodes und P.F. Stanbury, Hrsgb., IRL Press, S. 103-129; 131-163 und 165-192 10 (ISBN: 0199635773) und darin angegebenen Literaturstellen beschrieben.

Ein Beispiel ist die Analyse von Fettsäuren (Abkürzungen: FAME, Fettsäuremethylester; GC-MS, Gas-Flüssigkeitschromatographie-Mas- 15 senspektrometrie; TAG, Triacylglycerin; TLC, Dünnschichtchromatographie).

Der unzweideutige Nachweis für das Vorliegen von Fettsäureprodukten kann mittels Analyse rekombinanter Organismen nach Standard- 20 Analyseverfahren erhalten werden: GC, GC-MS oder TLC, wie verschiedentlich beschrieben von Christie und den Literaturstellen darin (1997, in: Advances on Lipid Methodology, Vierte Aufl.: Christie, Oily Press, Dundee, 119-169; 1998, Gaschromatographie-Massenspektrometrie-Verfahren, Lipide 33:343-353).

25 Das zu analysierende Material kann durch Ultraschallbehandlung, Mahlen in der Glasmühle, flüssigen Stickstoff und Mahlen oder über andere anwendbare Verfahren aufgebrochen werden. Das Material muss nach dem Aufbrechen zentrifugiert werden. Das Sediment 30 wird in Aqua dest. resuspendiert, 10 min bei 100°C erhitzt, auf Eis abgekühlt und erneut zentrifugiert, gefolgt von Extraktion in 0,5 M Schwefelsäure in Methanol mit 2 % Dimethoxypropan für 1 Std. bei 90°C, was zu hydrolysierten Öl- und Lipidverbindungen führt, die transmethylierte Lipide ergeben. Diese Fettsäuremethylester 35 werden in Petrolether extrahiert und schließlich einer GC-Analyse unter Verwendung einer Kapillarsäule (Chrompack, WCOT Fused Silica, CP-Wax-52 CB, 25 Mikrom, 0,32 mm) bei einem Temperaturgradienten zwischen 170°C und 240°C für 20 min und 5 min bei 240°C unterworfen. Die Identität der erhaltenen Fettsäuremethylester 40 muss unter Verwendung von Standards, die aus kommerziellen Quellen erhältlich sind (d.h. Sigma), definiert werden.

Für die Öl-Analyse der mit den Suppressionskonstrukten transformierten Arabidopsis Pflanzen wird folgendes Protokoll angewendet:

45



## 76

Die Extraktion der Lipide aus Samen wird nach der Methode von Bligh & Dyer (1959) Can J Biochem Physiol 37:911 durchgeführt. Dazu werden 5 mg Arabidopsis Samen in 1,2 ml Qiagen-Microtubes (Qiagen, Hilden) auf einer Sartorius (Göttingen) Mikrowaage abgewogen. Das Samenmaterial wird mit 500 uL Chloroform/Methanol (2:1; enthält Mono-C17-glycerin von Sigma als internen Standard) in der Rätשמühle MM300 der Firma Retsch (Haan) homogenisiert und 20 min bei RT inkubiert. Nach Zugabe von 500 uL 50 mM Kaliumphosphatpuffer pH 7,5 erfolgt die Phasentrennung. Von der organischen Phase werden 50 µL abgenommen, mit 1500 uL Chloroform verdünnt und 5 µL auf die Kapillaren Chromarods SIII der Firma Iatroscan (SKS, Bechenheim) aufgetragen. Nach Auftrag der Proben werden diese für 15 min in einer Dünnschichtkammer, die gesättigt ist mit 6:2:2 Chloroform: Methanol: Toluol in einem ersten Schritt aufgetrennt. Nach Ablauf der Zeit werden die Kapillaren 4 min bei Raumtemperatur getrocknet und dann für 22 min in eine Dünnschichtkammer, die gesättigt ist mit 7:3 n-Hexan:Diethylether gestellt. Nach einem weiteren Trocknungsschritt für 4 min bei Raumtemperatur werden die Proben in einem Iatroscan MK-5 (SKS, Bechenheim) entsprechend Fraser & Taggart, 1988 J. Chromatogr. 439:404 analysiert. Folgende Parameter wurden für die Messungen eingestellt: Slice width 50 msec, Treshold 20 mV, Noise 30, Skim ratio 0. Die Quantifizierung der Daten erfolgte anhand des internen Standards Mono-C17-glycerin (Sigma) sowie einer erstellten Eichkurve mit Tri-C17-glycerin (Sigma) mittels des Programms ChromStar (SKS, Beichenheim).

Für die quantitative Bestimmung der Ölgehalte werden Samen von jeweils 10 Pflanzen derselben unabhängigen transgenen Linie analysiert. Insgesamt wurde der Ölgehalt von 30 transgene Linien der T1 Generation, 10 transgene Linien mit je 10 Pflanzen der T2 Generation und 5 transgene Linien mit je 10 Pflanzen der T3 Linien bestimmt. Dabei zeigen die transgenen Pflanzen einen signifikant höheren Ölgehalt als entsprechend gleichbehandelte Kontrollpflanzen.

## Beispiel 7:

Zum Nachweis der Funktionalität der multiplen RNAi Konstrukte wurden Gene ausgewählt, deren Supression einen deutlichen phänologischen Effekt hervorrufen. Ein solches Gen ist zum Beispiel Toc159. Dieses Gen ist essentiell für die Entwicklung und Funktionalität von Chloroplasten in Arabidopsis (Bauer et al. Nature, 403, 203-207). Ein Ausschalten dieses Gens führt zu chlorophylldefizienten Pflanzen, deren Blatt-Erscheinungsbild dann hell-grün bis weiss ist. Dieser Albino-Phänotyp ist sehr leicht zu unterscheiden von normalen Pflanzen.

## 77

Als weiteres visuelles Reprotergen wurde GFP, das grün-fluoreszierende Protein aus der Qualle *Aequorea victoria* eingesetzt. Dieses Reprotergen ist ein häufig verwendetes Reprotergen in Pflanzen (siehe z.B. Stewart, Plant Cell Rep 2001 20(5):376-82).

- 5 Ausgehend von *Arabidopsis thaliana* cDNA oder vom Plasmid pEGFP (BD Clontech, Heidelberg, Genbank-Eintrag U476561) wurde über PCR mittels der aufgeführten Oligonukleotide erzeugt. Dabei wurde folgendes Protokoll eingesetzt:

- 10 Zusammensetzung des PCR-Ansatzes (50 µL):

5,00 µL Template cDNA oder genomische DNA (ca. 1 µg)  
 5,00 µL 10x Puffer (Advantage-Polymerase) + 25 mM MgCl<sub>2</sub>  
 5,00 µL 2mM dNTP  
 15 1,25 µL je Primer (10 pmol/µL)  
 0,50 µL Advantage-Polymerase (Clontech)

PCR-Programm: Anfangsdenaturierung für 2 min bei 95°C, dann 35 Zyklen mit 45 sec 95°C, 45 sec 55°C und 2 min 72°C. Abschliessende

- 20 Extension von 5 min bei 72°C.

a) Ausgangsvektor pGEM-Toc159: Ausgehend von *Arabidopsis* cDNA wurde mit nachfolgendem Oligonukleotid-Primerpaar ein Fragment aus Toc159 (Genbank Acc.-No. T14P8.24) amplifiziert:

25

ONP18 (SEQ ID NO: 123):  
 5'-CTCGAGGAATTCATGGACTCAAAGTCGGTTACTCCA

ONP19 (SEQ ID NO: 124):

- 30 5'-GGATCCATAAGCAAGCTTTCTCACTCTCCCCATCTGTGGA

Das PCR Produkt wurde in den Vektor pGEM-T easy von Promega (Mannheim) gemäss Herstellerangaben kloniert, resultierend in dem pGEM-Toc159 Vektor und die Sequenz überprüft.

35

b) Ausgangsvektor pGEM-GFP: Ausgehend von dem Plasmid pEGFP (BD Clontech, Heidelberg, GenbankAcc.-No.: U476561) wurde mit nachfolgendem Oligonukleotid-Primerpaar ein Fragment aus GFP amplifiziert:

40

ONP20 (SEQ ID NO: 125): 5'-AAGCTTCCAACACTTGTCCTACTTTT  
 ONP21 (SEQ ID NO: 126): 5'-GGATCCTTAAAGCTCATCATGTTTGT

Das PCR Produkt wurde in den Vektor pGEM-T easy von Promega

- 45 (Mannheim) gemäss Herstellerangaben kloniert, resultierend in dem pGEM-GFP Vektor und die Sequenz überprüft.

## 78

c) Herstellung des Konstruktes pGEM-159-GFP Der Vektor pGEM-GFP wurde mit den Restriktionsenzymen HindIII und BamHI für 2 Stunden inkubiert. Parallel wurde der Vektor pGEM-Toc159 mit den gleichen Restriktionsenzymen inkubiert, anschliessend dann zusätzlich für 15 min mit alkalischer Phosphatase behandelt. Die alkalische Phosphatase wurde anschliessend durch Erhitzen auf 95 °C für 10 min inaktiviert. Die entstandenen DNA-Fragmente aus beiden Ansätzen wurden über Agarose-Gelelektrophorese aufgetrennt. Das 558 bp Fragment aus pGEM-GFP sowie das 3471 bp Fragment von pGEM-Toc159 wurden aus dem Gel ausgeschnitten und mit dem „Gelpurification“-Kit (Qiagen) nach Herstellerangaben aufgereinigt. Beide Fragmente wurden für 2 h bei 16°C ligiert (T4 Ligase, New England Biolabs) und anschliessend nach Herstellerangaben in E. coli DH5α Zellen (Stratagen) transformiert. Positive Klone wurden durch PCR mit dem Primerpaar OPN1 und OPN4 identifiziert und anschliessend verifiziert durch Sequenzierung. Der erhaltene Vektor wurde mit pGEM-159-GFP bezeichnet.

d) Herstellung des Suppressionskonstruktes 1: Der Vektor pGEM-159-GFP wurde einerseits mit den Restriktionsenzymen XhoI und BamHI, ein weiterer Ansatz mit BamHI und SalI inkubiert. Der zweite Ansatz mit BamHI/ SalI wurde anschliessend für weitere 15 min mit alkalischer Phosphatase inkubiert. Die DNA-Fragmente aus beiden Ansätzen wurden über Agarose-Gelelektrophorese aufgetrennt und folgende Fragmente ausgeschnitten: Ansatz BamHI-XhoI das 1091 bp Fragment; Ansatz BamHI-SalI das 4029 bp Fragment. Beide Fragmente wurden nach Aufreinigung aus dem Agarose-Gel (siehe oben) für 2 h bei 16°C mit T4 Ligase inkubiert und anschliessend in E. coli DH5α Zellen (Stratagen) transformiert. Positive Klone wurden durch PCR mit dem Primerpaar OPN1 identifiziert und anschliessend verifiziert durch Sequenzierung. Der erhaltene Vektor wurde als Suppressionskonstrukt 1 bezeichnet.

e) Herstellung des Suppressionskonstruktes 2: Das Suppressionskonstrukt 1 und der Vektor p3300.1 (Andreas Hilbrunner, Dissertation ETH Zürich, 2003) wurden für 2h Stunden mit dem Restriktionsenzym EcoRI inkubiert. Anschliessend wurde der Vektor p3300.1 15 min mit alkalischer Phosphatase behandelt. Beide Ansätze wurden gemischt und für 2 h bei 16°C mit T4 Ligase inkubiert. Der Ligationsansatz wurde dann in E. coli DH5α Zellen (Stratagen) transformiert. Das entstandene Suppressionskonstrukt 2 wurde dann für die Agrobacterium- und Pflanzentransformation eingesetzt. Die Nukleinsäuresequenz kodierend für Suppressionskonstrukt 2 (p3300.1-Toc159-GFP-RNAi) ist unter SEQ ID NO: 122 wiedergegeben.

45

- Die Transformation von Agrobakterien und Pflanzen wurde wie in Beispiel 3 bzw. 4 beschrieben durchgeführt. Zum Nachweis der Funktionalität des Suppressionskonstruktes 2 wurde dieses durch die nach Bechtold et al., 1993 (C.R. Acad. Sci. Ser. III Sci. Vie., 316, 1194-1199) beschriebene Blüten-Transformationsmethode in Arabidopsis transformiert. Aus Ausgangsmaterial wurden Arabidopsis Pflanzen der Varietät Columbia-0 verwendet, die bereits die T-DNA des binären Vektors pBIN-35S-GFP enthielten.
- 10 Durch Anregung durch ultraviolettes Licht im Wellenlängenbereich 470-490 nm die grüne Fluoreszenz von GFP in diesen Pflanzen angeregt werden und damit die Expression des eingebrachten Transgens überprüft werden. Dazu wurden Keimlinge 1 Woche nach Keimung oder Blattstücke bei älteren Pflanzen mit dem Fluoreszenzmikroskop
- 15 MZFLIII von Leica analysiert. Zur Anregung von GFP wurden folgende Parameter eingestellt: Quecksilberlampe HBO 100W/DC, Filter GFP3, Bildbearbeitung Leica-Software. Speziell die Verwendung eines Filters (GFP3), der oberhalb einer Wellenlänge von 525 nm nicht mehr durchlässig ist, ermöglicht die GFP-Analyse von grünen
- 20 Blattmaterial. Ohne diesen Filter könnte die starke Autofluoreszenz des Blattfarbstoffes Chlorophyll nicht ausgeschlossen werden. Die zur Transformation verwendete Arabidopsis Linie zeigte eine starke GFP Expression nach mikroskopischer Analyse.
- 25 Transformierte Samen wurden direkt auf Erde ausgelegt und angezogen. Nach einer Woche wurde nach Keimlingen gesucht, die keinen oder einen reduzierten Anteil des Blattfarbstoffes Chlorophyll enthielten. Solche Pflanzen waren leicht an ihrer hellgrünen oder weissen Erscheinungsbild zu erkennen. Diese Pflanzen wurden dann
- 30 weiter durch Fluoreszenz-Mikroskopie untersucht und mit entsprechend parallel gewachsenen grünen Pflanzen verglichen. Fig.5A zeigt beispielhaft ein solche identifizierte Pflanze, die sich deutlich in der Farbe der Blätter von parallel gewachsenen Pflanzen unterscheidet. Dabei ist der Albino-Phänotyp (weisse Blätter)
- 35 auf die Wirkung des Toc159-Suppressionskonstruktes zurückzuführen. Die nicht transformierten Nachkommen der mit Agrobacterium-Suspension behandelten Pflanzen zeigen den Albino-Phänotyp nicht. Der auftretende Albino-Phänotyp ist damit ein spezifischer Effekt des eingebrachten Suppressionskonstruktes.
- 40 Die Fluoreszenz-mikroskopische Untersuchung der Albino-Pflanzen zeigte dann (Fig.5B), dass keine GFP-Signale in solchen Pflanzen gefunden werden konnte. Im Vergleich dazu zeigten die parallel gewachsenen grünen Pflanzen deutliche GFP Signale. Die Abwesen-
- 45 heit des GFP-Signals in allen identifizierten Albino-Pflanzen demonstriert die Funktionalität des Suppressionskonstruktes, denn nur die mit dem Suppressionkonstrukt transformierten Pflanzen

80

zeigten keine GFP-Signale mehr. Es konnte keine Segregation der beiden angestrebten Phänotypen beobachtet werden. Damit konnte gezeigt werden, dass durch Verwendung von nur einem Kontrollelement (Promotor) zwei funktionell völlig unterschiedliche Gene, 5 die ihrerseits durch unterschiedliche Kontrollelemente in ihrer Expression reguliert werden, ausgeschaltet werden konnten.

10

15

20

25

30

35

40

45

## Patentansprüche

1. Verfahren zur Verminderung der Expression von mindestens zwei  
5 verschiedenen, endogenen Zielgenen in einer eukaryotischen  
Zelle oder einem eukaryotischen Organismus durch Einbringen  
eines zumindest teilweise doppelsträngigen Ribonukleinsäure-  
moleküls in besagte eukaryotische Zelle oder besagten euka-  
ryotischen Organismus, wobei das doppelsträngige Ribonuklein-  
10 säuremolekül umfasst
- a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei je-  
weils mindestens eine dieser "sense"-Ribonukleotidsequen-  
zen im wesentlichen identisch ist zu mindestens einem  
15 Teil des "sense"-RNA-Transkriptes eines jeden der besag-  
ten endogenen Zielgene und
- b) "antisense"-Ribonukleotidsequenzen, die zu besagten  
"sense"-Ribonukleotidsequenzen unter a) im wesentlichen  
20 komplementären sind.
2. Verfahren nach Anspruch 1, wobei die transkribierten RNAs von  
mindestens zwei der in ihrer Expression verminderten Zielgene  
untereinander eine Homologie von unter 90% haben.  
25
3. Verfahren nach Anspruch 1 oder 2, wobei die doppelsträngige  
RNA durch ein einziges selbstkomplementäres Ribonukleotidmo-  
lekül gebildet wird.
- 30 4. Verfahren nach einem der Ansprüche 1 bis 3, wobei mindestens  
eine der ausgehend von den einzelnen "sense"-Ribonukleotidse-  
quenzen gebildeten doppelsträngigen RNA-Strukturen eine Länge  
eines geradzahligen Vielfachen von 21 oder 22 Basenpaaren  
hat.  
35
5. Verfahren nach einem der Ansprüche 1 bis 4, wobei das Ribonu-  
kleotidmolekül zwischen mindestens einer "sense"-Ribonukleo-  
tidsequenz und der dazu im wesentlichen komplementären "anti-  
sense"-Ribonukleotidsequenz eine Ribonukleotidsequenz kodie-  
40 rend für ein Intron enthält.

## 82

6. Verfahren nach einem der Ansprüche 1 bis 5, dadurch gekennzeichnet, dass mindestens zwei der endogenen Zielgene ausgewählt sind aus jeweils unterschiedlichen Klassen von Speicherprotein ausgewählt aus den Speicherprotein-Klassen der 2S-Albumine, 7S-Globuline, 11S/12S-Globuline oder Zein-Proline.
7. Verfahren nach einem der Ansprüche 1 bis 6, wobei mindestens eine "sense"-Ribonukleotidsequenz im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes
- a) einer Speicherprotein-Nukleinsäuresequenz gemäß SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 59, 61, 63, 65, 67, 69, 71, 93, 95, 97, 99, 101, 103, 105, 107, 109 oder 112, oder
- b) eines Gens aus dem Homogentisatabbauweg gemäß SEQ ID NO: 115, 116, 118 oder 120, oder
- c) eines Gens ausgewählt aus der Gruppe bestehend aus Acetyltransacylasen, Acyltransportproteinen, Fettsäuredesaturasen, Malonyltransacylasen,  $\beta$ -Ketoacyl-ACP-synthetasen, 3-Keto-ACP-reduktasen, Enoyl-ACP-hydrasen, Thioesterasen, Enoyl-ACP-reduktasen, ADP-Glucosepyrophosphorylasen, Phosphorylasen, Stärkesynthetasen, Q-Enzymen, Sucrose-6-phosphatsynthetasen, Sucrose-6-phosphatphosphatasen, ADP-Glucosepyrophosphorylasen, Branching-Enzymen, Debranching-Enzymen, Amylasen, Chalconsynthasen, Chalconisomerasen, Phenylalaninammonialyasen, Dehydrokaempferol(flavone)hydroxylasen, Dihydroflavonolreduktasen, Dihydroflavanol-2-hydroxylasen, Flavonoid-3'-hydroxylasen, Flavonoid-5'-hydroxylasen, Flavonoidglycosyltransferasen, Flavonoidmethyltransferasen, Flavonoidacyltransferasen, Polygalacturonasen, Cellulasen, Pectinesterasen,  $\beta$ -(1-4)Glucanasen,  $\beta$ -Galactanasen, 1-Aminocyclopropan-1-carboxylatsynthasen, Phytoendesaturasen, Cinnamoyl-CoA:NADPH-Reduktasen, Cinnamoylalkoholdehydrogenasen, Coffeinsäure-O-methyltransferasen, Cinnamoylalkoholdehydrogenasen, Polyphenoloxidasen, Homogentisat-1,2-dioxygenasen, Maleylacetoacetatisomerasen, Fumarylacetoacetathydrolasen, N-Methyl-putrescinoxidasen, Putrescin-N-methyltransferasen, 7-Methylxanthine-3-methyltransferasen, 1-Methylxanthin-3-methyltransferasen und Threoninsynthasen.

## 83

8. Ribonukleinsäuremolekül, das eine zumindest teilweise doppelsträngige Struktur hat und umfasst
- a) mindestens zwei "sense"-Ribonukleotidsequenzen, wobei jeweils mindestens eine dieser "sense"-Ribonukleotidsequenzen im wesentlichen identisch ist zu mindestens einem Teil des "sense"-RNA-Transkriptes eines endogenen Zielgens, wobei jedoch nicht alle "sense"-Ribonukleotidsequenzen zu dem "sense"-RNA-Transkript eines einzigen endogenen Zielgens im wesentlichen identisch sind, und
- b) "antisense"-Ribonukleotidsequenzen, die zu besagten "sense"-Ribonukleotidsequenzen unter a) im wesentlichen komplementären sind.
9. Ribonukleinsäuremolekül nach Anspruch 8, wobei das Ribonukleinsäuremolekül wie in einem der Ansprüche 2 bis 7 gekennzeichnet ist.
10. Transgene Expressionskassette enthaltend in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für doppelsträngiges Ribonukleinsäuremolekül gemäß einem der Ansprüche 8 oder 9, wobei das Ribonukleinsäuremolekül aus einem einzigen RNA-Strang gebildet wird.
11. Transgenes Expressionssystem enthaltend
- a) in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für "sense"-Ribonukleotidsequenzen eines doppelsträngigen Ribonukleinsäuremoleküls gemäß einem der Ansprüche 8 oder 9 und
- b) in funktioneller Verknüpfung mit einem Promotor eine Nukleinsäuresequenz kodierend für "antisense"-Ribonukleotidsequenzen eines doppelsträngigen Ribonukleinsäuremoleküls gemäß einem der Ansprüche 8 oder 9,
- wobei das Ribonukleinsäuremolekül aus den beiden unter a) und b) definierten Strängen gebildet wird, und die Promotoren so gewählt sind, das in einem bestimmten Organismus oder Zelle die gleichzeitige Expression von "sense"-Ribonukleotidsequenzen und "antisense"-Ribonukleotidsequenzen gewährleistet ist.



## 84

12. Transgener Vektor enthaltend eine transgene Expressionskassette gemäß Anspruch 10 oder ein transgenes Expressionssystem gemäß Anspruch 11.
- 5 13. Transgener Organismus enthaltend eine transgene Expressionskassette gemäß Anspruch 10 oder ein transgenes Expressionssystem gemäß Anspruch 11 oder einen transgenen Vektor gemäß Anspruch 12.
- 10 14. Transgener Organismus nach Anspruch 13 ausgewählt aus der Gruppe bestehend aus Bakterien, Hefen, nicht-menschlichen Tieren und Pflanzen.
- 15 15. Transgener Organismus nach Ansprüchen 13, dadurch gekennzeichnet, dass die Pflanze ausgewählt ist aus der Gruppe der landwirtschaftlichen Nutzpflanzen.
- 20 16. Verwendung eines Ribonukleotidmoleküls nach einem der Ansprüche 8 oder 9, einer transgenen Expressionskassette gemäß Anspruch 10, eines transgenen Expressionssystems gemäß Anspruch 11, eines transgenen Vektors gemäß Anspruch 12 oder eines transgenen Organismus gemäß einem der Ansprüche 13 bis 15 zur Herstellung von Arzneimitteln, in biotechnologischen Verfahren oder in der Pflanzenbiotechnologie.
- 25 17. Verwendung nach Anspruch 16, wobei mindestens einer der nachfolgenden Eigenschaften in Pflanzen erzielt wird:
- 30 a) Verbesserter Schutz gegen abiotische Stressfaktoren
- b) Modifikation der Zusammensetzung und/oder des Gehaltes an Fettsäuren, Lipiden oder Ölen
- 35 c) Modifikation der Kohlenhydratzusammensetzung
- d) Veränderung der Farbe oder Pigmentierung
- e) Verminderung des Gehaltes von Speicherproteinen
- 40 f) Erreichen einer Resistenz gegen pflanzliche Pathogene
- g) Verhinderung von Halmbruch
- h) Verzögerung der Fruchtreifung
- 45 i) Erzielen einer männlichen Sterilität

## 85

- j) Verminderung unerwünschter oder toxischer Pflanzeninhaltsstoffe
- 5 k) Verzögerung von Alterserscheinungen
- l) Modifikation der Lignifikation und/oder des Ligningehaltes
- 10 m) Modifikation des Faseranteils in Nahrungsmitteln oder der Faserqualität in Baumwolle
- n) Verminderung der Stoßanfälligkeit
- 15 o) Steigerung der Vitamin E Biosynthese
- p) Verminderung des Nikotingehaltes, des Coffeingehaltes oder des Theophyllin-Gehaltes
- 20 q) Erhöhung des Methioningehaltes durch Verminderung der Threoninbiosynthese
- 25
- 30
- 35
- 40
- 45

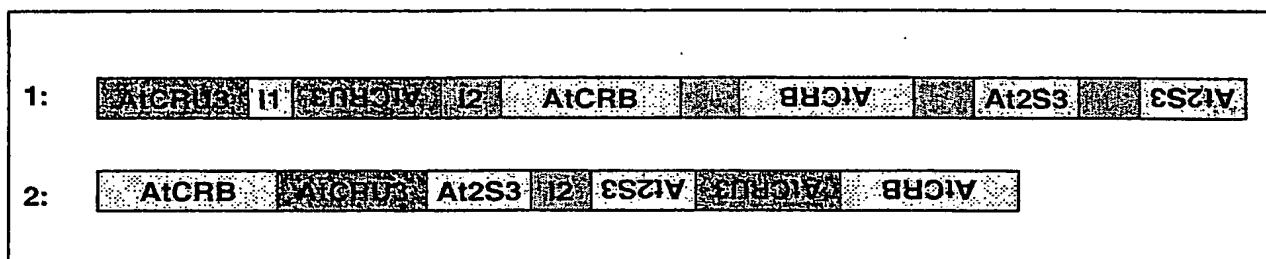


Fig. 1

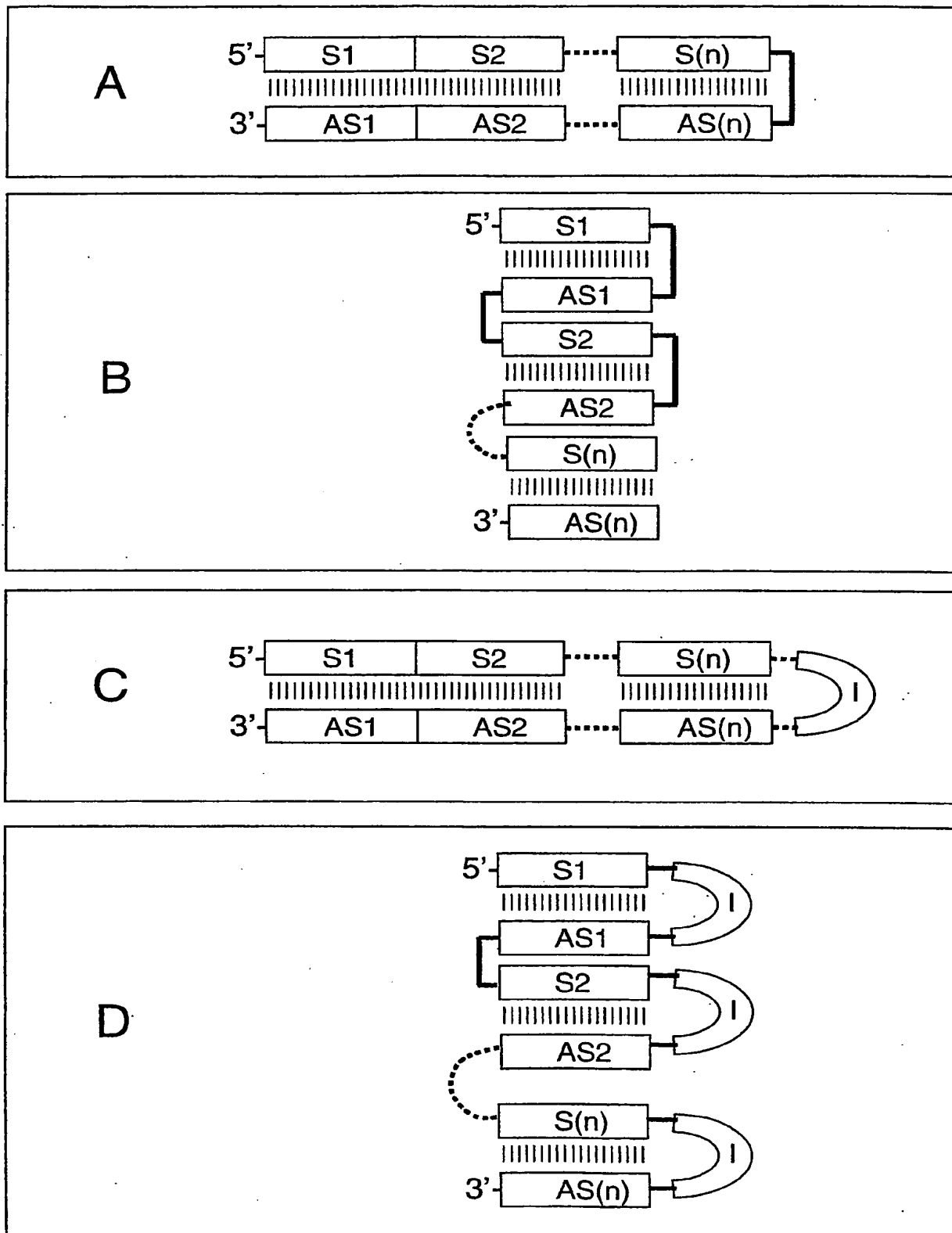


Fig.2

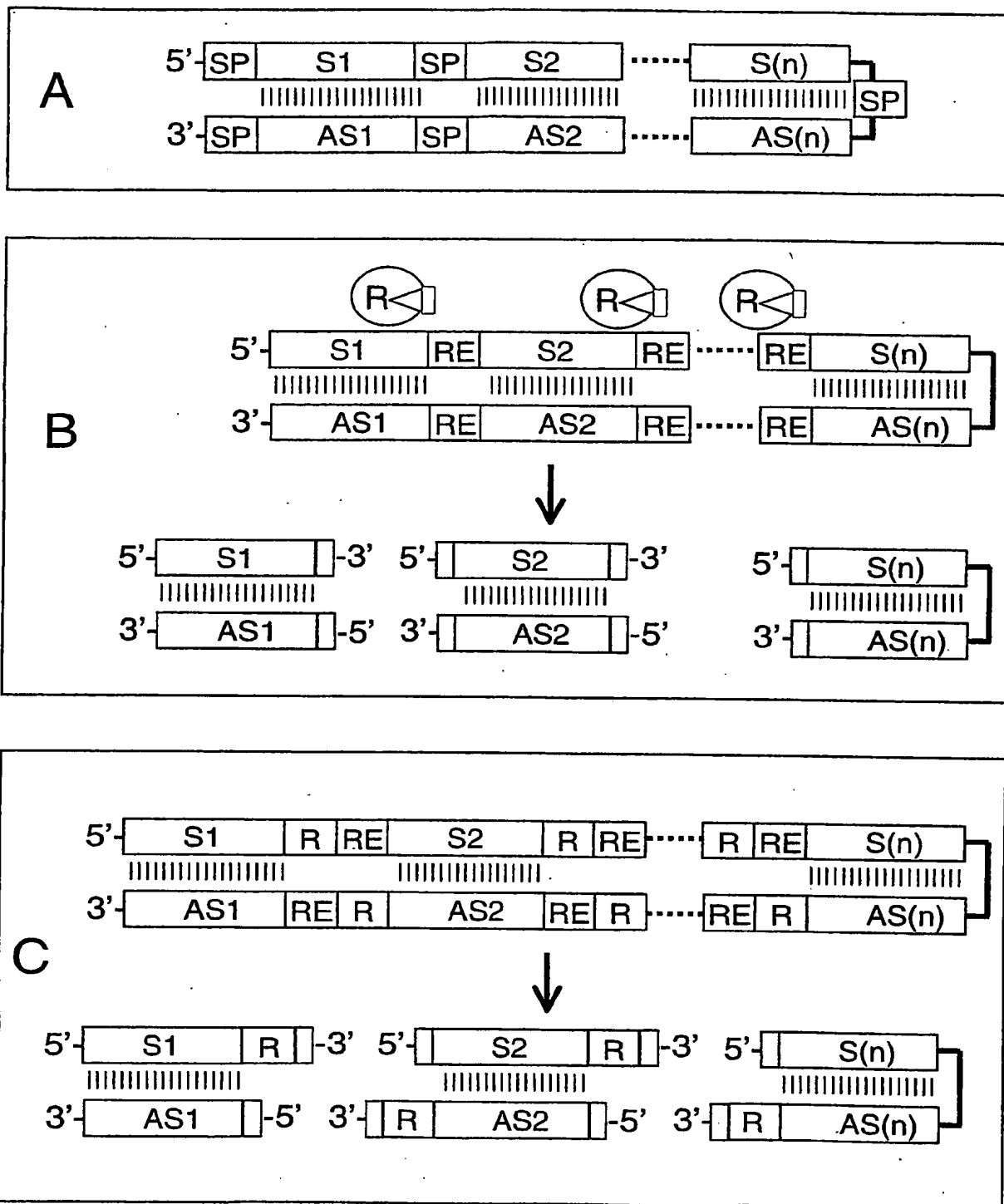


Fig.3

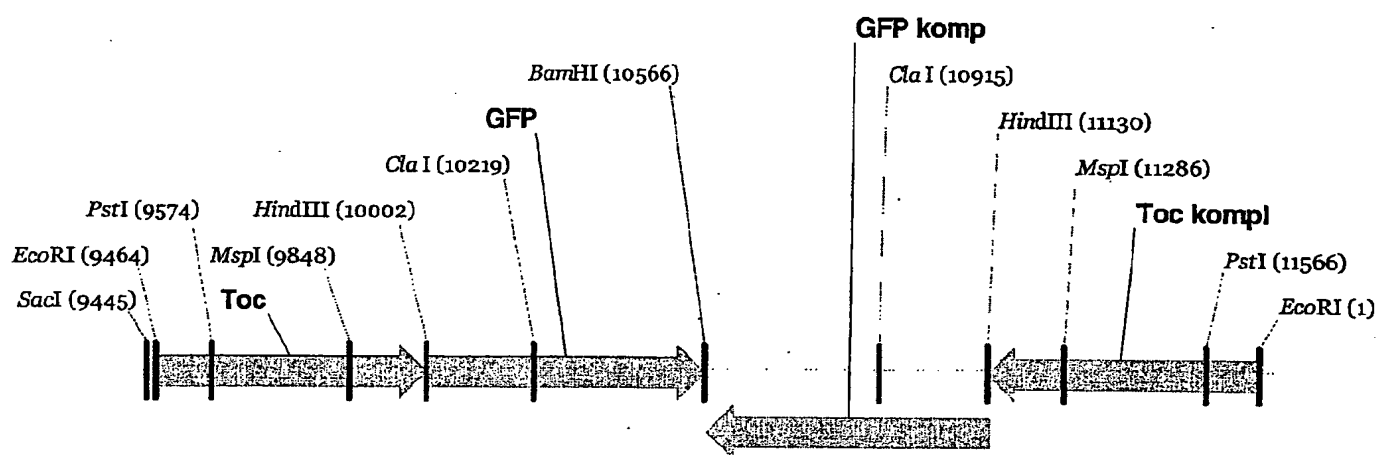
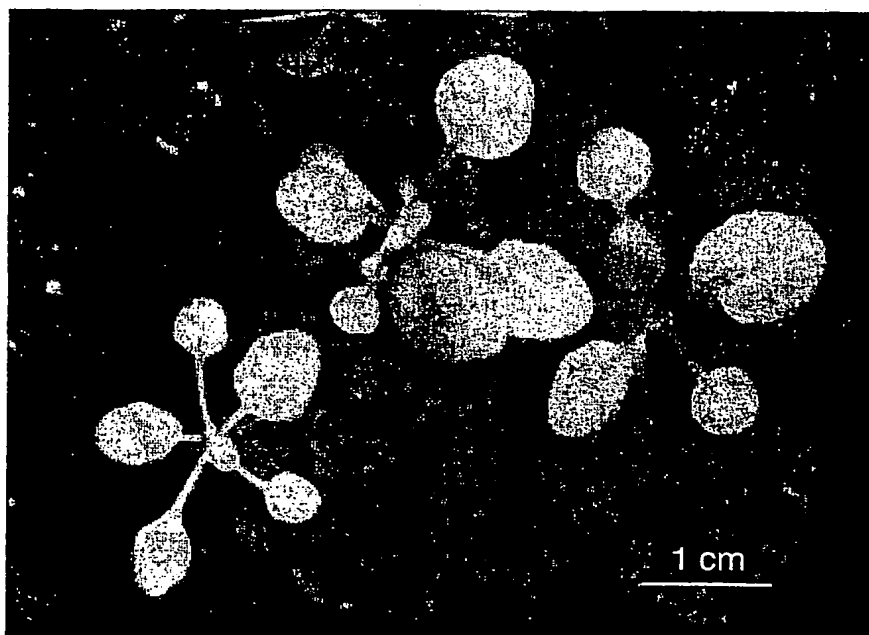


Fig.4

A



B

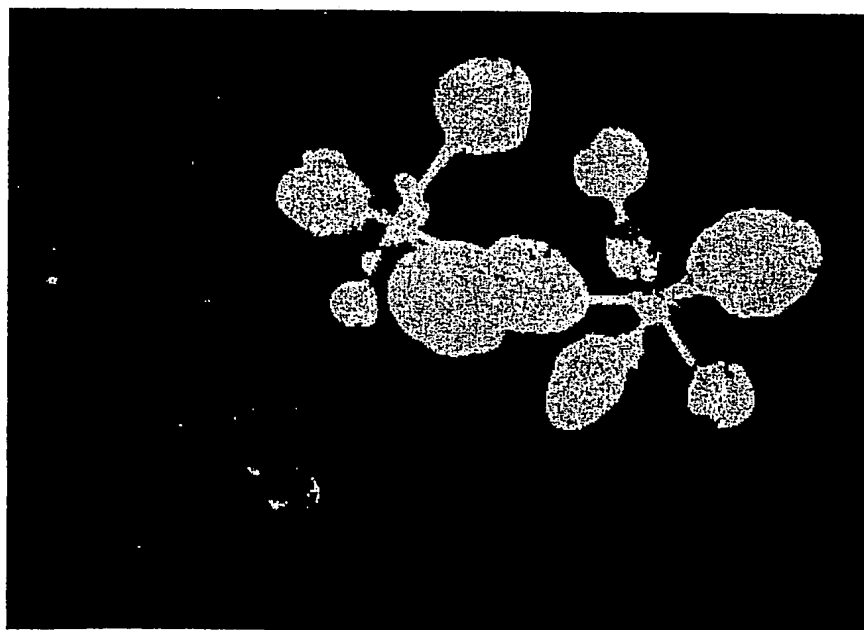


Fig.5

## SEQUENZPROTOKOLL

&lt;110&gt; BASF Plant Science GmbH

&lt;120&gt; Konstrukte und Verfahren zur Regulation der Genexpression

&lt;130&gt; PD009300062-AT

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 126

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 495

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(492)

&lt;223&gt; albumine 2S subunit 1

&lt;400&gt; 1

atg gca aac aag ttg ttc ctc gtc tgc gca gct ctc gct ctc tgc ttc	48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe	
1 5 10 15	
ctc ctc acc aac gct tcc atc tac cgc acc gtc gtt gag ttc gaa gaa	96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu	
20 25 30	
gat gac gcc act aac ccc ata ggc cca aaa atg agg aaa tgc cgc aag	144
Asp Asp Ala Thr Asn Pro Ile Gly Pro Lys Met Arg Lys Cys Arg Lys	
35 40 45	
gag ttt cag aaa gaa caa cac cta aga gct tgc cag caa ttg atg ctc	192
Glu Phe Gln Lys Glu Gln His Leu Arg Ala Cys Gln Gln Leu Met Leu	
50 55 60	
cag caa gca agg caa ggc cgt agc gat gag ttt gat ttc gaa gac gac	240
Gln Gln Ala Arg Gln Gly Arg Ser Asp Glu Phe Asp Phe Glu Asp Asp	
65 70 75 80	
atg gag aac cca cag gga caa cag cag gaa caa cag cta ttc cag cag	288
Met Glu Asn Pro Gln Gly Gln Gln Gln Glu Gln Gln Leu Phe Gln Gln	
85 90 95	
tgc tgc aac gag ctt cgc cag gaa gag cca gat tgt gtt tgc ccc acc	336
Cys Cys Asn Glu Leu Arg Gln Glu Glu Pro Asp Cys Val Cys Pro Thr	
100 105 110	
ttg aaa caa gct gcc aag gcc gtt aga ctc cag gga cag cac caa cca	384
Leu Lys Gln Ala Ala Lys Ala Val Arg Leu Gln Gly Gln His Gln Pro	
115 120 125	
atg caa gtc agg aaa att tac cag aca gcc aag cac ttg ccc aac gtt	432
Met Gln Val Arg Lys Ile Tyr Gln Thr Ala Lys His Leu Pro Asn Val	
130 135 140	
tgc gac atc ccg caa gtt gat gtt tgt ccc ttc aac atc cct tca ttc	480
Cys Asp Ile Pro Gln Val Asp Val Cys Pro Phe Asn Ile Pro Ser Phe	
145 150 155 160	
cct tct ttc tac taa	495
Pro Ser Phe Tyr	



<210> 2  
 <211> 164  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 2  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe  
 1 5 10 15  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu  
 20 25 30  
 Asp Asp Ala Thr Asn Pro Ile Gly Pro Lys Met Arg Lys Cys Arg Lys  
 35 40 45  
 Glu Phe Gln Lys Glu Gln His Leu Arg Ala Cys Gln Gln Leu Met Leu  
 50 55 60  
 Gln Gln Ala Arg Gln Gly Arg Ser Asp Glu Phe Asp Phe Glu Asp Asp  
 65 70 75 80  
 Met Glu Asn Pro Gln Gly Gln Gln Gln Glu Gln Gln Leu Phe Gln Gln  
 85 90 95  
 Cys Cys Asn Glu Leu Arg Gln Glu Glu Pro Asp Cys Val Cys Pro Thr  
 100 105 110  
 Leu Lys Gln Ala Ala Lys Ala Val Arg Leu Gln Gly Gln His Gln Pro  
 115 120 125  
 Met Gln Val Arg Lys Ile Tyr Gln Thr Ala Lys His Leu Pro Asn Val  
 130 135 140  
 Cys Asp Ile Pro Gln Val Asp Val Cys Pro Phe Asn Ile Pro Ser Phe  
 145 150 155 160  
 Pro Ser Phe Tyr

<210> 3  
 <211> 495  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (1)..(492)  
 <223> albumine 2S subunit 3  
 <400> 3  
 atg gct aac aag ctc ttc ctc gtc tgc gca act ctc gcc ctc tgc ttc 48  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Leu Ala Leu Cys Phe  
 1 5 10 15  
 ctc ctc acc aac gct tcc atc tac cgc acc gtt gtc gaa ttc gaa gaa 96  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu  
 20 25 30  
 gat gac gcc agc aac ccc gta ggt cca aga cag aga tgc cag aag gag 144  
 Asp Asp Ala Ser Asn Pro Val Gly Pro Arg Gln Arg Cys Gln Lys Glu  
 35 40 45  
 ttt cag caa tca caa cac cta aga gct tgc cag aga tgg atg agc aag 192  
 Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Arg Trp Met Ser Lys  
 50 55 60

## 3

caa atg agg caa gga cgt ggt ggt ggt cct tcc ctc gac gat gag ttc 240  
 Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe  
 65 70 75 80  
 gat ttc gag ggc ccc cag cag gga tac cag cta ctc cag cag tgc tgc 288  
 Asp Phe Glu Gly Pro Gln Gln Gly Tyr Gln Leu Leu Gln Gln Cys Cys  
 85 90 95  
 aac gag ctt cgc cag gaa gag cca gtt tgc gtt tgc ccc acc ttg aaa 336  
 Asn Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro Thr Leu Lys  
 100 105 110  
 caa gct gcc agg gca gtt agc ctc cag gga cag cac gga cca ttc caa 384  
 Gln Ala Ala Arg Ala Val Ser Leu Gln Gly Gln His Gly Pro Phe Gln  
 115 120 125  
 tcc agg aaa att tac cag tca gct aag tac ttg cct aac att tgc aag 432  
 Ser Arg Lys Ile Tyr Gln Ser Ala Lys Tyr Leu Pro Asn Ile Cys Lys  
 130 135 140  
 atc cag caa gtt ggt gaa tgt ccc ttc cag acc acc atc cct ttc ttc 480  
 Ile Gln Gln Val Gly Glu Cys Pro Phe Gln Thr Thr Ile Pro Phe Phe  
 145 150 155 160  
 cct cct tac tac tag 495  
 Pro Pro Tyr Tyr  
 <210> 4  
 <211> 164  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 4  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Leu Ala Leu Cys Phe  
 1 5 10 15  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Glu Glu  
 20 25 30  
 Asp Asp Ala Ser Asn Pro Val Gly Pro Arg Gln Arg Cys Gln Lys Glu  
 35 40 45  
 Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Arg Trp Met Ser Lys  
 50 55 60  
 Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe  
 65 70 75 80  
 Asp Phe Glu Gly Pro Gln Gln Gly Tyr Gln Leu Leu Gln Gln Cys Cys  
 85 90 95  
 Asn Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro Thr Leu Lys  
 100 105 110  
 Gln Ala Ala Arg Ala Val Ser Leu Gln Gly Gln His Gly Pro Phe Gln  
 115 120 125  
 Ser Arg Lys Ile Tyr Gln Ser Ala Lys Tyr Leu Pro Asn Ile Cys Lys  
 130 135 140  
 Ile Gln Gln Val Gly Glu Cys Pro Phe Gln Thr Thr Ile Pro Phe Phe  
 145 150 155 160  
 Pro Pro Tyr Tyr

&lt;210&gt; 5

&lt;211&gt; 513

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(510)

&lt;223&gt; albumine 2S subunit 2

&lt;400&gt; 5

```

atg gca aac aag ctc ttc ctc gtc tgc gca act ttc gcc ctc tgc ttc      48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Phe Ala Leu Cys Phe
   1               5               10               15

ctc ctc acc aac gct tcc atc tac cgc act gtt gtc gag ttc gac gaa      96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
               20               25               30

gat gac gcc agc aac ccc atg ggc cca aga cag aaa tgt cag aag gag     144
Asp Asp Ala Ser Asn Pro Met Gly Pro Arg Gln Lys Cys Gln Lys Glu
               35               40               45

ttt cag caa tca cag cac cta aga gct tgc cag aaa ttg atg cgc atg     192
Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Lys Leu Met Arg Met
               50               55               60

caa atg agg caa ggc cgt ggt ggt ggt ccc tcc ctc gac gat gag ttc     240
Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
               65               70               75               80

gat ttg gaa gac gac atc gag aac cca caa ggc ccc cag cag gga cac     288
Asp Leu Glu Asp Asp Ile Glu Asn Pro Gln Gly Pro Gln Gln Gly His
               85               90               95

cag atc ctc cag cag tgc tgc agc gag ctt cgc cag gaa gag cca gtt     336
Gln Ile Leu Gln Gln Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val
               100              105              110

tgt gtt tgc ccc acc ttg aga caa gct gcc agg gcc gtt agc ctc cag     384
Cys Val Cys Pro Thr Leu Arg Gln Ala Ala Arg Ala Val Ser Leu Gln
               115              120              125

gga caa cac gga cca ttc caa tcc agg aaa att tac aag aca gct aag     432
Gly Gln His Gly Pro Phe Gln Ser Arg Lys Ile Tyr Lys Thr Ala Lys
               130              135              140

tac ttg cct aac att tgc aag atc cag caa gtt ggt gaa tgc ccc ttc     480
Tyr Leu Pro Asn Ile Cys Lys Ile Gln Gln Val Gly Glu Cys Pro Phe
               145              150              155              160

cag acc acc atc cct ttc ttc cct cct tac taa                        513
Gln Thr Thr Ile Pro Phe Phe Pro Pro Tyr
               165              170

```

&lt;210&gt; 6

&lt;211&gt; 170

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 6

```

Met Ala Asn Lys Leu Phe Leu Val Cys Ala Thr Phe Ala Leu Cys Phe
   1               5               10               15

Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
               20               25               30

Asp Asp Ala Ser Asn Pro Met Gly Pro Arg Gln Lys Cys Gln Lys Glu
               35               40               45

```

Phe Gln Gln Ser Gln His Leu Arg Ala Cys Gln Lys Leu Met Arg Met  
 50 55 60  
 Gln Met Arg Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe  
 65 70 75 80  
 Asp Leu Glu Asp Asp Ile Glu Asn Pro Gln Gly Pro Gln Gln Gly His  
 85 90 95  
 Gln Ile Leu Gln Gln Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val  
 100 105 110  
 Cys Val Cys Pro Thr Leu Arg Gln Ala Ala Arg Ala Val Ser Leu Gln  
 115 120 125  
 Gly Gln His Gly Pro Phe Gln Ser Arg Lys Ile Tyr Lys Thr Ala Lys  
 130 135 140  
 Tyr Leu Pro Asn Ile Cys Lys Ile Gln Gln Val Gly Glu Cys Pro Phe  
 145 150 155 160  
 Gln Thr Thr Ile Pro Phe Phe Pro Pro Tyr  
 165 170

&lt;210&gt; 7

&lt;211&gt; 501

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(498)

&lt;223&gt; albumine 2S subunit 4

&lt;400&gt; 7

atg gcg aac aag ctc ttc ctc gtc tgc gca gct ctc gcc ctg tgt ttc 48  
 Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe  
 1 5 10 15  
 atc ctc acc aac gct tcc gtc tat cgc acc gtt gtc gag ttc gac gaa 96  
 Ile Leu Thr Asn Ala Ser Val Tyr Arg Thr Val Val Glu Phe Asp Glu  
 20 25 30  
 gat gac gcc agt aac ccc ata ggc cca ata cag aaa tgt cag aag gag 144  
 Asp Asp Ala Ser Asn Pro Ile Gly Pro Ile Gln Lys Cys Gln Lys Glu  
 35 40 45  
 ttt cag caa gac cag cac cta aga gct tgc cag aga tgg atg cgc aag 192  
 Phe Gln Gln Asp Gln His Leu Arg Ala Cys Gln Arg Trp Met Arg Lys  
 50 55 60  
 caa atg tgg caa gga cgt ggt ggt ggt cct tcc ctc gac gat gag ttc 240  
 Gln Met Trp Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe  
 65 70 75 80  
 gat atg gaa gac gac atc gag aac ccg cag aga cga cag cta ctc cag 288  
 Asp Met Glu Asp Asp Ile Glu Asn Pro Gln Arg Arg Gln Leu Leu Gln  
 85 90 95  
 aag tgc tgc agc gag ctt cgc caa gaa gag cca gtt tgc gtt tgc ccc 336  
 Lys Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro  
 100 105 110  
 acc ttg aga caa gct gcc aag gcc gtt aga ttc cag gga cag caa cac 384  
 Thr Leu Arg Gln Ala Ala Lys Ala Val Arg Phe Gln Gly Gln Gln His  
 115 120 125

```

caa cca gag caa gtc agg aaa att tac cag gca gct aag tac ttg cct 432
Gln Pro Glu Gln Val Arg Lys Ile Tyr Gln Ala Ala Lys Tyr Leu Pro
    130                135                140

aac att tgc aaa atc cag caa gtt ggt gtt tgc ccc ttc cag atc cct 480
Asn Ile Cys Lys Ile Gln Gln Val Gly Val Cys Pro Phe Gln Ile Pro
    145                150                155                160

tca atc cct tct tac tac taa
Ser Ile Pro Ser Tyr Tyr
                165

```

&lt;210&gt; 8

&lt;211&gt; 166

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 8

```

Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Leu Ala Leu Cys Phe
  1                5                10                15

Ile Leu Thr Asn Ala Ser Val Tyr Arg Thr Val Val Glu Phe Asp Glu
                20                25                30

Asp Asp Ala Ser Asn Pro Ile Gly Pro Ile Gln Lys Cys Gln Lys Glu
    35                40                45

Phe Gln Gln Asp Gln His Leu Arg Ala Cys Gln Arg Trp Met Arg Lys
    50                55                60

Gln Met Trp Gln Gly Arg Gly Gly Gly Pro Ser Leu Asp Asp Glu Phe
    65                70                75                80

Asp Met Glu Asp Asp Ile Glu Asn Pro Gln Arg Arg Gln Leu Leu Gln
                85                90                95

Lys Cys Cys Ser Glu Leu Arg Gln Glu Glu Pro Val Cys Val Cys Pro
    100                105                110

Thr Leu Arg Gln Ala Ala Lys Ala Val Arg Phe Gln Gly Gln Gln His
    115                120                125

Gln Pro Glu Gln Val Arg Lys Ile Tyr Gln Ala Ala Lys Tyr Leu Pro
    130                135                140

Asn Ile Cys Lys Ile Gln Gln Val Gly Val Cys Pro Phe Gln Ile Pro
    145                150                155                160

Ser Ile Pro Ser Tyr Tyr
                165

```

&lt;210&gt; 9

&lt;211&gt; 1473

&lt;212&gt; DNA

&lt;213&gt; Brassica napus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1470)

&lt;223&gt; cruciferin

&lt;400&gt; 9

```

atg gct cgg ctc tca tct ctt ctc tct ttt tcc tta gca ctt ttg atc 48
Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Ile
  1                5                10                15

```

ttt ctc cat ggc tct aca gct caa cag ttt cca aac gag tgt cag cta	96
Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu	
20 25 30	
gac cag ctc aat gca ctg gag ccg tca cac gta ctt aag gct gag gct	144
Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala	
35 40 45	
ggg cgc atc gag gtg tgg gac cac cac gct cct cag cta cgt tgc tct	192
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser	
50 55 60	
ggg gtc tcc ttt gta cgt tac atc atc gag tct aag ggt ctc tac ttg	240
Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu	
65 70 75 80	
ccc tct ttc ttt agc acc gcg aag ctc tcc ttc gtg gct aaa gga gaa	288
Pro Ser Phe Phe Ser Thr Ala Lys Leu Ser Phe Val Ala Lys Gly Glu	
85 90 95	
ggg ctt atg ggg aga gtg gtc cct gga tgc gcc gag aca ttc cag gac	336
Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp	
100 105 110	
tca tca gtg ttt caa cca agc ggt ggt agc ccc tcg gga gaa ggt cag	384
Ser Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Ser Gly Glu Gly Gln	
115 120 125	
ggc caa gga caa caa ggt cag ggc caa ggc cac caa ggt caa ggc caa	432
Gly Gln Gly Gln Gln Gly Gln Gly Gln Gly His Gln Gly Gln Gly Gln	
130 135 140	
gga caa cag ggc caa caa ggt cag caa gga caa cag agt caa ggc cag	480
Gly Gln Gln Gly Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln	
145 150 155 160	
ggc ttc cgt gat atg cac cag aaa gtg gag cac ata agg act ggg gac	528
Gly Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp	
165 170 175	
acc atc gct aca cat ccc ggt gta gcc caa tgg ttc tac aac gac gga	576
Thr Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly	
180 185 190	
aac caa cca ctt gtc atc gtt tcc gtc ctc gat tta gcc agc cac cag	624
Asn Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln	
195 200 205	
aat cag ctc gac cgc aac cca agg cca ttt tac tta gcc gga aac aac	672
Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn	
210 215 220	
cca caa ggc caa gta tgg ata gaa gga cgc gag caa cag cca caa aag	720
Pro Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys	
225 230 235 240	
aac atc ctt aat ggc ttc aca cca gag gtt ctt gct aaa gct ttc aag	768
Asn Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys	
245 250 255	
atc gat gtt agg aca gcg caa caa ctt cag aac cag caa gac aac cgt	816
Ile Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asn Arg	
260 265 270	
gga aac att atc cga gtc caa ggc cca ttc agt gtc att agg ccg cct	864
Gly Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro	
275 280 285	

```

ttg agg agt cag aga ccg cag gag aca gaa gtt aac ggt tta gaa gag 912
Leu Arg Ser Gln Arg Pro Gln Glu Thr Glu Val Asn Gly Leu Glu Glu
290 295 300

acc ata tgc agc gcg agg tgc acc gat aac ctc gat gac cca tct aat 960
Thr Ile Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn
305 310 315 320

gct gac gta tac aag cca cag ctc ggt tac atc agc act ctg aac agc 1008
Ala Asp Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser
325 330 335

tat gat ctc ccc atc ctt cgc ttc ctt cgt ctc tca gcc ctc cgt gga 1056
Tyr Asp Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly
340 345 350

tct atc cgt caa aac gcg atg gtg ctt cca cag tgg aac gca aac gca 1104
Ser Ile Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala
355 360 365

aac gcg gtt ctc tac gtg aca gac ggg gaa gcc cat gtg cag gtg gtt 1152
Asn Ala Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val
370 375 380

aac gac aac ggt gac aga gtg ttc gac gga caa gtc tct caa gga cag 1200
Asn Asp Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln
385 390 395 400

cta ctt tcc ata cca caa ggt ttc tcc gtg gtg aaa cgc gca aca agc 1248
Leu Leu Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser
405 410 415

gaa cag ttc cgg tgg atc gag ttc aag aca aac gca aac gca cag atc 1296
Glu Gln Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile
420 425 430

aac aca ctt gct gga cga acc tcg gtc ttg aga ggt tta cca tta gag 1344
Asn Thr Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu
435 440 445

gtc ata tcc aat ggg tac caa atc tca ctc gaa gaa gca aga agg gtt 1392
Val Ile Ser Asn Gly Tyr Gln Ile Ser Leu Glu Glu Ala Arg Arg Val
450 455 460

aag ttc aac acg atc gag acc act ttg acg cac agc agt ggc cca gct 1440
Lys Phe Asn Thr Ile Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala
465 470 475 480

agc tac gga ggg cca agg aag gct gat gct taa 1473
Ser Tyr Gly Gly Pro Arg Lys Ala Asp Ala
485 490

<210> 10
<211> 490
<212> PRT
<213> Brassica napus

<400> 10
Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Ile
1 5 10 15
Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu
20 25 30
Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala
35 40 45

```

Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser  
 50 55 60  
 Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu  
 65 70 75 80  
 Pro Ser Phe Phe Ser Thr Ala Lys Leu Ser Phe Val Ala Lys Gly Glu  
 85 90 95  
 Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp  
 100 105 110  
 Ser Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Ser Gly Glu Gly Gln  
 115 120 125  
 Gly Gln Gly Gln Gln Gly Gln Gly Gln Gly His Gln Gly Gln Gly Gln  
 130 135 140  
 Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln  
 145 150 155 160  
 Gly Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp  
 165 170 175  
 Thr Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly  
 180 185 190  
 Asn Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln  
 195 200 205  
 Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn  
 210 215 220  
 Pro Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys  
 225 230 235 240  
 Asn Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys  
 245 250 255  
 Ile Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg  
 260 265 270  
 Gly Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro  
 275 280 285  
 Leu Arg Ser Gln Arg Pro Gln Glu Thr Glu Val Asn Gly Leu Glu Glu  
 290 295 300  
 Thr Ile Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn  
 305 310 315 320  
 Ala Asp Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser  
 325 330 335  
 Tyr Asp Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly  
 340 345 350  
 Ser Ile Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala  
 355 360 365  
 Asn Ala Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val  
 370 375 380  
 Asn Asp Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln  
 385 390 395 400  
 Leu Leu Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser  
 405 410 415  
 Glu Gln Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile  
 420 425 430



Asn	Thr	Leu	Ala	Gly	Arg	Thr	Ser	Val	Leu	Arg	Gly	Leu	Pro	Leu	Glu
		435					440					445			
Val	Ile	Ser	Asn	Gly	Tyr	Gln	Ile	Ser	Leu	Glu	Glu	Ala	Arg	Arg	Val
	450					455					460				
Lys	Phe	Asn	Thr	Ile	Glu	Thr	Thr	Leu	Thr	His	Ser	Ser	Gly	Pro	Ala
465					470					475					480
Ser	Tyr	Gly	Gly	Pro	Arg	Lys	Ala	Asp	Ala						
				485					490						

&lt;210&gt; 11

&lt;211&gt; 1467

&lt;212&gt; DNA

&lt;213&gt; Brassica napus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1464)

&lt;223&gt; cruciferin

&lt;400&gt; 11

atg	gct	cgg	ctc	tca	tct	ctt	ctc	tct	ttt	tcc	tta	gca	ctt	ttg	act	48
Met	Ala	Arg	Leu	Ser	Ser	Leu	Leu	Ser	Phe	Ser	Leu	Ala	Leu	Leu	Thr	
1				5					10					15		
ttt	ctc	cat	ggc	tct	aca	gct	caa	cag	ttt	cca	aac	gag	tgt	cag	cta	96
Phe	Leu	His	Gly	Ser	Thr	Ala	Gln	Gln	Phe	Pro	Asn	Glu	Cys	Gln	Leu	
			20					25					30			
gac	cag	ctc	aat	gca	ctg	gag	ccg	tca	cac	gta	ctt	aag	gct	gag	gct	144
Asp	Gln	Leu	Asn	Ala	Leu	Glu	Pro	Ser	His	Val	Leu	Lys	Ala	Glu	Ala	
			35				40					45				
ggg	cgc	atc	gag	gtg	tgg	gac	cac	cac	gct	cct	cag	cta	cgt	tgc	tct	192
Gly	Arg	Ile	Glu	Val	Trp	Asp	His	His	Ala	Pro	Gln	Leu	Arg	Cys	Ser	
	50					55					60					
ggg	gtc	tcc	ttt	gta	cgt	tac	atc	atc	gag	tct	aag	ggg	ctc	tac	ttg	240
Gly	Val	Ser	Phe	Val	Arg	Tyr	Ile	Ile	Glu	Ser	Lys	Gly	Leu	Tyr	Leu	
65					70				75						80	
ccc	tct	ttc	ttt	agc	acc	gcg	agg	ctc	tcc	ttc	gtg	gct	aaa	gga	gaa	288
Pro	Ser	Phe	Phe	Ser	Thr	Ala	Arg	Leu	Ser	Phe	Val	Ala	Lys	Gly	Glu	
				85					90					95		
ggg	ctt	atg	ggg	aga	gtg	gtc	ctg	tgc	gcc	gag	aca	ttc	cag	gac	tca	336
Gly	Leu	Met	Gly	Arg	Val	Val	Leu	Cys	Ala	Glu	Thr	Phe	Gln	Asp	Ser	
			100					105					110			
tca	gtg	ttt	caa	cca	agc	ggg	ggg	agc	ccc	ttc	gga	gaa	ggg	cag	ggc	384
Ser	Val	Phe	Gln	Pro	Ser	Gly	Gly	Ser	Pro	Phe	Gly	Glu	Gly	Gln	Gly	
		115					120					125				
caa	gga	caa	caa	ggg	cag	ggc	caa	ggc	cac	caa	ggg	caa	ggc	caa	gga	432
Gln	Gly	Gln	Gln	Gly	Gln	Gly	Gln	Gly	His	Gln	Gly	Gln	Gly	Gln	Gly	
	130					135					140					
caa	cag	ggc	caa	caa	ggg	cag	caa	gga	caa	cag	agt	caa	ggc	cag	ggg	480
Gln	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Ser	Gln	Gly	Gln	Gly	
145					150				155						160	
ttc	cgt	gat	atg	cac	cag	aaa	gtg	gag	cac	ata	agg	act	ggg	gac	acc	528
Phe	Arg	Asp	Met	His	Gln	Lys	Val	Glu	His	Ile	Arg	Thr	Gly	Asp	Thr	
				165					170					175		

atc gct aca cat ccc ggt gta gcc caa tgg ttc tac aac gac gga aac	576
Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly Asn	
180 185 190	
caa cca ctt gtc atc gtt tcc gtc ctc gat tta gcc agc cac cag aat	624
Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln Asn	
195 200 205	
cag ctc gac cgc aac cca agg cca ttt tac tta gcc gga aac aac cca	672
Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn Pro	
210 215 220	
caa ggc caa gta tgg ata gaa gga cgc gag caa cag cca caa aag aac	720
Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys Asn	
225 230 235 240	
atc ctt aat ggc ttc aca cca gag gtt ctt gct aaa gct ttc aag atc	768
Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys Ile	
245 250 255	
gat gtt agg aca gcg caa caa ctt cag aac cag caa gac aac cgt gga	816
Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg Gly	
260 265 270	
aac att atc cga gtc caa ggc cca ttc agt gtc att agg ccg cct ttg	864
Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro Leu	
275 280 285	
agg agt cag aga ccg cag gag gaa gtt aac ggt tta gaa gag acc ata	912
Arg Ser Gln Arg Pro Gln Glu Glu Val Asn Gly Leu Glu Glu Thr Ile	
290 295 300	
tgc agc gcg agg tgc acc gat aac ctc gat gac cca tct aat gct gac	960
Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn Ala Asp	
305 310 315 320	
gta tac aag cca cag ctc ggt tac atc agc act ctg aac agc tat gat	1008
Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp	
325 330 335	
ctc ccc atc ctt cgc ttc ctt cgt ctc tca gcc ctc cgt gga tct atc	1056
Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly Ser Ile	
340 345 350	
cgt caa aac gcg atg gtg ctt cca cag tgg aac gca aac gca aac gcg	1104
Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala	
355 360 365	
gtt ctc tac gtg aca gac ggg gaa gcc cat gtg cag gtg gtt aac gac	1152
Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val Asn Asp	
370 375 380	
aac ggt gac aga gtg ttc gac gga caa gtc tct caa gga cag cta ctt	1200
Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Leu	
385 390 395 400	
tcc ata cca caa ggt ttc tcc gtg gtg aaa cgc gca aca agc gaa cag	1248
Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Glu Gln	
405 410 415	
ttc cgg tgg atc gag ttc aag aca aac gca aac gca cag atc aac aca	1296
Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr	
420 425 430	
ctt gct gga cga acc tcg gtc ttg aga ggt tta cca tta gag gtc ata	1344
Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile	
435 440 445	

tcc aat ggg tac caa atc tca ctc gaa gaa gca aga agg gtt aag ttc 1392  
 Ser Asn Gly Tyr Gln Ile Ser Leu Glu Glu Ala Arg Arg Val Lys Phe  
 450 455 460

aac acg atc gag acc act ttg acg cac agc agt ggc cca gct agc tac 1440  
 Asn Thr Ile Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr  
 465 470 475 480

gga ggg cca agg aag gct gat gct taa 1467  
 Gly Gly Pro Arg Lys Ala Asp Ala  
 485

<210> 12

<211> 488

<212> PRT

<213> Brassica napus

<400> 12

Met Ala Arg Leu Ser Ser Leu Leu Ser Phe Ser Leu Ala Leu Leu Thr  
 1 5 10 15

Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu  
 20 25 30

Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala  
 35 40 45

Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser  
 50 55 60

Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Lys Gly Leu Tyr Leu  
 65 70 75 80

Pro Ser Phe Phe Ser Thr Ala Arg Leu Ser Phe Val Ala Lys Gly Glu  
 85 90 95

Gly Leu Met Gly Arg Val Val Leu Cys Ala Glu Thr Phe Gln Asp Ser  
 100 105 110

Ser Val Phe Gln Pro Ser Gly Gly Ser Pro Phe Gly Glu Gly Gln Gly  
 115 120 125

Gln Gly Gln Gln Gly Gln Gly Gln Gly His Gln Gly Gln Gly Gln Gly  
 130 135 140

Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Ser Gln Gly Gln Gly  
 145 150 155 160

Phe Arg Asp Met His Gln Lys Val Glu His Ile Arg Thr Gly Asp Thr  
 165 170 175

Ile Ala Thr His Pro Gly Val Ala Gln Trp Phe Tyr Asn Asp Gly Asn  
 180 185 190

Gln Pro Leu Val Ile Val Ser Val Leu Asp Leu Ala Ser His Gln Asn  
 195 200 205

Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr Leu Ala Gly Asn Asn Pro  
 210 215 220

Gln Gly Gln Val Trp Ile Glu Gly Arg Glu Gln Gln Pro Gln Lys Asn  
 225 230 235 240

Ile Leu Asn Gly Phe Thr Pro Glu Val Leu Ala Lys Ala Phe Lys Ile  
 245 250 255

Asp Val Arg Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg Gly  
 260 265 270

Asn Ile Ile Arg Val Gln Gly Pro Phe Ser Val Ile Arg Pro Pro Leu  
 275 280 285  
 Arg Ser Gln Arg Pro Gln Glu Glu Val Asn Gly Leu Glu Glu Thr Ile  
 290 295 300  
 Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Asn Ala Asp  
 305 310 315 320  
 Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp  
 325 330 335  
 Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly Ser Ile  
 340 345 350  
 Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala  
 355 360 365  
 Val Leu Tyr Val Thr Asp Gly Glu Ala His Val Gln Val Val Asn Asp  
 370 375 380  
 Asn Gly Asp Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Leu  
 385 390 395 400  
 Ser Ile Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Glu Gln  
 405 410 415  
 Phe Arg Trp Ile Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr  
 420 425 430  
 Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile  
 435 440 445  
 Ser Asn Gly Tyr Gln Ile Ser Leu Glu Glu Ala Arg Arg Val Lys Phe  
 450 455 460  
 Asn Thr Ile Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr  
 465 470 475 480  
 Gly Gly Pro Arg Lys Ala Asp Ala  
 485

&lt;210&gt; 13

&lt;211&gt; 1491

&lt;212&gt; DNA

&lt;213&gt; Brassica napus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1488)

&lt;223&gt; cruciferin BnC2

&lt;400&gt; 13

atg gct cga ctc tcg tct ctt ctc tat ttt tcg ata aca gtt ttg atc	48
Met Ala Arg Leu Ser Ser Leu Leu Tyr Phe Ser Ile Thr Val Leu Ile	
1 5 10 15	
ttt ctc cat ggc tct aca gct caa cag ttt cca aac gag tgc caa cta	96
Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu	
20 25 30	
gac cag ctc aat gcg ctg gag ccg tca cac gta ctt aag gcc gag gct	144
Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala	
35 40 45	
ggt cgc atc gaa gtg tgg gac cac cac gct cct cag cta cgc tgc tct	192
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser	
50 55 60	

ggt gtc tcc ttc gta cgt tac ata atc gag tct cag ggt cta tac ttg	240
Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Gln Gly Leu Tyr Leu	
65 70 75 80	
ccc tct ttc tta aat acc gcg aac gtc tct ttc gtt gct aaa gga caa	288
Pro Ser Phe Leu Asn Thr Ala Asn Val Ser Phe Val Ala Lys Gly Gln	
85 90 95	
ggc ctt atg ggg aga gtg gtc cct gga tgc gct gag act ttc cag gac	336
Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp	
100 105 110	
tca tca gta ttc caa cca ggc agt ggc agc ccc ttc gga gaa ggt caa	384
Ser Ser Val Phe Gln Pro Gly Ser Gly Ser Pro Phe Gly Glu Gly Gln	
115 120 125	
ggc caa ggt cag cag ggt cag ggc caa ggt cag ggt cag ggt caa ggc	432
Gly Gln Gly Gln Gln Gly Gln Gly Gln Gly Gln Gly Gln Gly Gln Gly	
130 135 140	
aag ggc caa cag ggt caa ggc aag ggc caa cag ggt caa tcc cag ggc	480
Lys Gly Gln Gln Gly Gln Gly Lys Gly Gln Gln Gly Gln Ser Gln Gly	
145 150 155 160	
caa cag ggt caa ggt caa ggt ttc cgt gat atg cac cag aaa gta gag	528
Gln Gln Gly Gln Gly Gln Gly Phe Arg Asp Met His Gln Lys Val Glu	
165 170 175	
cac ata agg agc ggc gac acc att gct aca cat ccc ggt gta gct caa	576
His Ile Arg Ser Gly Asp Thr Ile Ala Thr His Pro Gly Val Ala Gln	
180 185 190	
tgg ttc tac aac aat gga aac caa cct ctt gtc atc gtt gcc gtc atg	624
Trp Phe Tyr Asn Asn Gly Asn Gln Pro Leu Val Ile Val Ala Val Met	
195 200 205	
gat tta gct agc cac cag aac cag ctt gac cgc aac cca agc caa ttt	672
Asp Leu Ala Ser His Gln Asn Gln Leu Asp Arg Asn Pro Ser Gln Phe	
210 215 220	
tac tta gca gga aaa aac cca caa ggc caa tca tgg cta cac gga cga	720
Tyr Leu Ala Gly Lys Asn Pro Gln Gly Gln Ser Trp Leu His Gly Arg	
225 230 235 240	
ggg caa cag cca caa aac aac atc ctt aat ggc ttc tct cca gag gtt	768
Gly Gln Gln Pro Gln Asn Asn Ile Leu Asn Gly Phe Ser Pro Glu Val	
245 250 255	
ctt gct caa gcg ttc aag atc gat gtt agg aca gcg caa caa ctt cag	816
Leu Ala Gln Ala Phe Lys Ile Asp Val Arg Thr Ala Gln Gln Leu Gln	
260 265 270	
aac cag caa gat aac cgg gga aac att gtc cgt gtc caa ggc ccc ttc	864
Asn Gln Gln Asp Asn Arg Gly Asn Ile Val Arg Val Gln Gly Pro Phe	
275 280 285	
ggt gtt att agg ccg cca ttg aaa agc cag aga cca cag gag aca gaa	912
Gly Val Ile Arg Pro Pro Leu Lys Ser Gln Arg Pro Gln Glu Thr Glu	
290 295 300	
gct aac ggt cta gaa gag acc ata tgc agc gca agg tgc acg gat aac	960
Ala Asn Gly Leu Glu Glu Thr Ile Cys Ser Ala Arg Cys Thr Asp Asn	
305 310 315 320	
ctc gat gac cca tct aac gcg gat gtg tat aag cca cag ctt ggt tac	1008
Leu Asp Asp Pro Ser Asn Ala Asp Val Tyr Lys Pro Gln Leu Gly Tyr	
325 330 335	

## 15

atc agc att ctt aac agt tat gat cta ccc atc ctt cgc gta ctt cgc	1056
Ile Ser Ile Leu Asn Ser Tyr Asp Leu Pro Ile Leu Arg Val Leu Arg	
340 345 350	
ctc tca gcc ctc cgt gga tca atc cgt caa aat gca atg gtt ctt cca	1104
Leu Ser Ala Leu Arg Gly Ser Ile Arg Gln Asn Ala Met Val Leu Pro	
355 360 365	
cag tgg aag tca aag tca aac gcg gtt ctc tac gtg aca gac ggg gaa	1152
Gln Trp Lys Ser Lys Ser Asn Ala Val Leu Tyr Val Thr Asp Gly Glu	
370 375 380	
gcc caa ata cag gtg gtt aac gac aac ggt gac aga gtg ttc gat gga	1200
Ala Gln Ile Gln Val Val Asn Asp Asn Gly Asp Arg Val Phe Asp Gly	
385 390 395 400	
caa gtc tct caa ggg cag cta ctt tcc att cca caa gga ttc tcc gtt	1248
Gln Val Ser Gln Gly Gln Leu Leu Ser Ile Pro Gln Gly Phe Ser Val	
405 410 415	
gtg aaa cgc gca aca agc gat cag ttc agg tgg ata gaa ttc aag aca	1296
Val Lys Arg Ala Thr Ser Asp Gln Phe Arg Trp Ile Glu Phe Lys Thr	
420 425 430	
aac gca aac gcc cag atc aac act ctt gct gga cgt acc tca gtc atg	1344
Asn Ala Asn Ala Gln Ile Asn Thr Leu Ala Gly Arg Thr Ser Val Met	
435 440 445	
aga ggt tta cca tta gag gtc ata gcc aat ggg tac caa atc tca ctt	1392
Arg Gly Leu Pro Leu Glu Val Ile Ala Asn Gly Tyr Gln Ile Ser Leu	
450 455 460	
gaa gaa gca aga agg gtt aag ttc aac aca ata gag acc act ttg acc	1440
Glu Glu Ala Arg Arg Val Lys Phe Asn Thr Ile Glu Thr Thr Leu Thr	
465 470 475 480	
cac agt agt ggc cca gcg agc tac gga agg cca agg aag gct gat gct	1488
His Ser Ser Gly Pro Ala Ser Tyr Gly Arg Pro Arg Lys Ala Asp Ala	
485 490 495	
tga	1491
<210> 14	
<211> 496	
<212> PRT	
<213> Brassica napus	
<400> 14	
Met Ala Arg Leu Ser Ser Leu Leu Tyr Phe Ser Ile Thr Val Leu Ile	
1 5 10 15	
Phe Leu His Gly Ser Thr Ala Gln Gln Phe Pro Asn Glu Cys Gln Leu	
20 25 30	
Asp Gln Leu Asn Ala Leu Glu Pro Ser His Val Leu Lys Ala Glu Ala	
35 40 45	
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser	
50 55 60	
Gly Val Ser Phe Val Arg Tyr Ile Ile Glu Ser Gln Gly Leu Tyr Leu	
65 70 75 80	
Pro Ser Phe Leu Asn Thr Ala Asn Val Ser Phe Val Ala Lys Gly Gln	
85 90 95	
Gly Leu Met Gly Arg Val Val Pro Gly Cys Ala Glu Thr Phe Gln Asp	
100 105 110	



<210> 15  
 <211> 555  
 <212> DNA  
 <213> Brassica napus  
 <220>  
 <221> CDS  
 <222> (1)..(552)  
 <223> cruciferin cru4  
 <400> 15  
 ttg tgc aca atg aga tgc acc gaa aac ctt gat gac ccg tca agt gct 48  
 Leu Cys Thr Met Arg Cys Thr Glu Asn Leu Asp Asp Pro Ser Ser Ala  
 1 5 10 15  
 gat gtc tac aag cca tcg ctc gga tac att agc aca ctc aac agc tac 96  
 Asp Val Tyr Lys Pro Ser Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr  
 20 25 30  
 aac ctc cct atc ctc aga ttc ctc cgc ctt agc gct ctt cgt ggc tcc 144  
 Asn Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly Ser  
 35 40 45  
 atc cat aac aac gct atg gtg ctg ccg caa tgg aac gtg aac gca aac 192  
 Ile His Asn Asn Ala Met Val Leu Pro Gln Trp Asn Val Asn Ala Asn  
 50 55 60  
 gcg gca ctc tac gtg aca aag ggg aag gct cat ata cag atg gtg aac 240  
 Ala Ala Leu Tyr Val Thr Lys Gly Lys Ala His Ile Gln Met Val Asn  
 65 70 75 80  
 gac aac gga caa aga gtg ttt gac caa gag atc tcc cag gga cag tta 288  
 Asp Asn Gly Gln Arg Val Phe Asp Gln Glu Ile Ser Gln Gly Gln Leu  
 85 90 95  
 ctt gtc gtg cca caa ggc ttc gcg gtc gtg aaa cgt gcc aca agc caa 336  
 Leu Val Val Pro Gln Gly Phe Ala Val Val Lys Arg Ala Thr Ser Gln  
 100 105 110  
 cag ttc cag tgg atc gag ttc aag agc aac gac aac gca cag atc aac 384  
 Gln Phe Gln Trp Ile Glu Phe Lys Ser Asn Asp Asn Ala Gln Ile Asn  
 115 120 125  
 aca ctc gcg gga cgc acc tca gtc atg aga ggt tta cca ctt gag gtt 432  
 Thr Leu Ala Gly Arg Thr Ser Val Met Arg Gly Leu Pro Leu Glu Val  
 130 135 140  
 ata tcc aac ggg tat cag atc tca ccc caa gaa gct aga agt gtt aag 480  
 Ile Ser Asn Gly Tyr Gln Ile Ser Pro Gln Glu Ala Arg Ser Val Lys  
 145 150 155 160  
 ttc agc act ctt gag acc aca ttg act caa agc agt ggt cct atg ggc 528  
 Phe Ser Thr Leu Glu Thr Thr Leu Thr Gln Ser Ser Gly Pro Met Gly  
 165 170 175  
 tac ggt atg cct aga gtc gag gct tga 555  
 Tyr Gly Met Pro Arg Val Glu Ala  
 180  
 <210> 16  
 <211> 184  
 <212> PRT  
 <213> Brassica napus  
 <400> 16  
 Leu Cys Thr Met Arg Cys Thr Glu Asn Leu Asp Asp Pro Ser Ser Ala  
 1 5 10 15



Asp Val Tyr Lys Pro Ser Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr  
                   20                  25                  30  
 Asn Leu Pro Ile Leu Arg Phe Leu Arg Leu Ser Ala Leu Arg Gly Ser  
                   35                  40                  45  
 Ile His Asn Asn Ala Met Val Leu Pro Gln Trp Asn Val Asn Ala Asn  
                   50                  55                  60  
 Ala Ala Leu Tyr Val Thr Lys Gly Lys Ala His Ile Gln Met Val Asn  
                   65                  70                  75                  80  
 Asp Asn Gly Gln Arg Val Phe Asp Gln Glu Ile Ser Gln Gly Gln Leu  
                   85                  90                  95  
 Leu Val Val Pro Gln Gly Phe Ala Val Val Lys Arg Ala Thr Ser Gln  
                   100                  105                  110  
 Gln Phe Gln Trp Ile Glu Phe Lys Ser Asn Asp Asn Ala Gln Ile Asn  
                   115                  120                  125  
 Thr Leu Ala Gly Arg Thr Ser Val Met Arg Gly Leu Pro Leu Glu Val  
                   130                  135                  140  
 Ile Ser Asn Gly Tyr Gln Ile Ser Pro Gln Glu Ala Arg Ser Val Lys  
                   145                  150                  155                  160  
 Phe Ser Thr Leu Glu Thr Thr Leu Thr Gln Ser Ser Gly Pro Met Gly  
                   165                  170                  175  
 Tyr Gly Met Pro Arg Val Glu Ala  
                   180

&lt;210&gt; 17

&lt;211&gt; 1530

&lt;212&gt; DNA

&lt;213&gt; Brassica napus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1527)

&lt;223&gt; cruciferin cru4

&lt;400&gt; 17

atg gtt aaa gtt cct cat ctc ctc gtc gca acg ttc ggg gtt ctc ctc 48  
 Met Val Lys Val Pro His Leu Leu Val Ala Thr Phe Gly Val Leu Leu  
   1                  5                  10                  15  
 gtc ctc aac ggc tgt ctc gca agg cag tcg cta ggg gtt cct cct cag 96  
 Val Leu Asn Gly Cys Leu Ala Arg Gln Ser Leu Gly Val Pro Pro Gln  
                   20                  25                  30  
 cta ggg aac gcg tgt aac ctc gat aac tta gac gtt ctc cag cct acc 144  
 Leu Gly Asn Ala Cys Asn Leu Asp Asn Leu Asp Val Leu Gln Pro Thr  
                   35                  40                  45  
 gaa act atc aag agc gag gct ggt cgg gtc gag tac tgg gat cac aac 192  
 Glu Thr Ile Lys Ser Glu Ala Gly Arg Val Glu Tyr Trp Asp His Asn  
                   50                  55                  60  
 aat cct cag atc cga tgt gct ggt gtc tct gtc tct cgt gtt ata atc 240  
 Asn Pro Gln Ile Arg Cys Ala Gly Val Ser Val Ser Arg Val Ile Ile  
                   65                  70                  75                  80  
 gaa caa ggc ggt ctc tac ctt cct acc ttc ttc agc tcc ccc aaa att 288  
 Glu Gln Gly Gly Leu Tyr Leu Pro Thr Phe Phe Ser Ser Pro Lys Ile  
                   85                  90                  95

tca	tac	gtt	gtt	caa	gga	atg	ggt	att	agc	gga	aga	gtg	gtc	cct	gga	336
Ser	Tyr	Val	Val	Gln	Gly	Met	Gly	Ile	Ser	Gly	Arg	Val	Val	Pro	Gly	
		100						105					110			
tgc	gcg	gaa	acc	ttc	atg	gac	tcg	cag	cct	atg	caa	gga	caa	caa	caa	384
Cys	Ala	Glu	Thr	Phe	Met	Asp	Ser	Gln	Pro	Met	Gln	Gly	Gln	Gln	Gln	
		115					120					125				
ggt	caa	cca	tgg	cag	gga	caa	caa	gga	caa	cag	ggt	cag	cag	gga	caa	432
Gly	Gln	Pro	Trp	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Gln	
		130				135					140					
caa	ggt	caa	cag	ggt	cag	cag	gga	caa	caa	ggt	caa	cag	ggt	cag	cag	480
Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Gln	Gln	
145					150					155					160	
ggt	caa	cag	gga	cag	cag	ggt	cag	cag	cag	caa	ggg	ttc	cgt	gac	atg	528
Gly	Gln	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gln	Gln	Gly	Phe	Arg	Asp	Met	
				165				170						175		
cac	cag	aag	gtc	gaa	cat	gtt	cga	cat	gga	gac	atc	att	gcc	att	act	576
His	Gln	Lys	Val	Glu	His	Val	Arg	His	Gly	Asp	Ile	Ile	Ala	Ile	Thr	
			180					185					190			
gca	ggc	tct	tcc	cat	tgg	atc	tac	aac	acc	ggt	gac	cag	cca	ctt	gtc	624
Ala	Gly	Ser	Ser	His	Trp	Ile	Tyr	Asn	Thr	Gly	Asp	Gln	Pro	Leu	Val	
		195					200					205				
att	atc	tgc	ctt	ctc	gac	att	gcc	aac	tac	caa	aac	caa	ctc	gac	cgc	672
Ile	Ile	Cys	Leu	Leu	Asp	Ile	Ala	Asn	Tyr	Gln	Asn	Gln	Leu	Asp	Arg	
		210				215					220					
aac	cca	aga	acg	ttc	cgt	ctg	gcc	gga	aac	aac	cca	cag	ggc	ggt	tcc	720
Asn	Pro	Arg	Thr	Phe	Arg	Leu	Ala	Gly	Asn	Asn	Pro	Gln	Gly	Gly	Ser	
225					230					235					240	
cag	cag	cag	cag	caa	caa	caa	cag	aac	atg	ttg	agc	ggg	ttc	gac	cct	768
Gln	Gln	Gln	Gln	Gln	Gln	Gln	Gln	Asn	Met	Leu	Ser	Gly	Phe	Asp	Pro	
				245				250						255		
cag	gtc	cta	gcc	cag	gca	ttg	aaa	atc	gac	gtt	agg	ttg	gct	cag	gag	816
Gln	Val	Leu	Ala	Gln	Ala	Leu	Lys	Ile	Asp	Val	Arg	Leu	Ala	Gln	Glu	
			260					265					270			
ctt	cag	aac	caa	caa	gac	agc	aga	gga	aac	atc	gtt	cgt	gtt	aag	gga	864
Leu	Gln	Asn	Gln	Gln	Asp	Ser	Arg	Gly	Asn	Ile	Val	Arg	Val	Lys	Gly	
		275					280					285				
cct	ttc	cag	gtt	gtg	agg	ccg	cct	ctt	aga	cag	cca	tac	gag	agt	gag	912
Pro	Phe	Gln	Val	Val	Arg	Pro	Pro	Leu	Arg	Gln	Pro	Tyr	Glu	Ser	Glu	
		290				295					300					
cag	tgg	aga	cac	ccc	cgt	ggc	cca	cca	caa	agc	cca	caa	gac	aac	ggc	960
Gln	Trp	Arg	His	Pro	Arg	Gly	Pro	Pro	Gln	Ser	Pro	Gln	Asp	Asn	Gly	
305					310					315					320	
ttg	gag	gag	act	atc	tgc	agc	atg	agg	acc	cac	gag	aac	att	gat	gac	1008
Leu	Glu	Glu	Thr	Ile	Cys	Ser	Met	Arg	Thr	His	Glu	Asn	Ile	Asp	Asp	
				325					330					335		
cca	gcc	cgt	gct	gac	gtg	tat	aag	ccc	aac	ctc	ggc	cgt	gtg	act	agc	1056
Pro	Ala	Arg	Ala	Asp	Val	Tyr	Lys	Pro	Asn	Leu	Gly	Arg	Val	Thr	Ser	
			340					345					350			
gtc	aac	agc	tac	act	tta	ccc	atc	ttg	cag	tat	atc	aga	ctc	agc	gcc	1104
Val	Asn	Ser	Tyr	Thr	Leu	Pro	Ile	Leu	Gln	Tyr	Ile	Arg	Leu	Ser	Ala	
		355					360					365				

acc cgt ggc att ctc cag ggt aat gcg atg gtg ctt ccg aaa tac aac 1152  
 Thr Arg Gly Ile Leu Gln Gly Asn Ala Met Val Leu Pro Lys Tyr Asn  
 370 375 380

atg aac gcg aac gag atc ttg tac tgc act caa gga caa gca agg att 1200  
 Met Asn Ala Asn Glu Ile Leu Tyr Cys Thr Gln Gly Gln Ala Arg Ile  
 385 390 395 400

caa gtg gtg aac gac aac gga cag aac gtg ctg gac cag cag gtg cag 1248  
 Gln Val Val Asn Asp Asn Gly Gln Asn Val Leu Asp Gln Gln Val Gln  
 405 410 415

aag gga cag ctc gtg gtc atc cca caa gga ttc gcc tat gtt gtc cag 1296  
 Lys Gly Gln Leu Val Val Ile Pro Gln Gly Phe Ala Tyr Val Val Gln  
 420 425 430

tcc cac caa aac aac ttc gaa tgg att tct ttc aag aca aac gct aac 1344  
 Ser His Gln Asn Asn Phe Glu Trp Ile Ser Phe Lys Thr Asn Ala Asn  
 435 440 445

gcg atg gtc agc act ttg gcc ggt aga acc tcg gcc ttg agg gca ttg 1392  
 Ala Met Val Ser Thr Leu Ala Gly Arg Thr Ser Ala Leu Arg Ala Leu  
 450 455 460

cca cta gag gtc ata acc aac gct ttc caa att tct ctc gag gaa gct 1440  
 Pro Leu Glu Val Ile Thr Asn Ala Phe Gln Ile Ser Leu Glu Glu Ala  
 465 470 475 480

aga agg atc aag ttc aac acg ctt gag acc act ttg act cgt gcg cgc 1488  
 Arg Arg Ile Lys Phe Asn Thr Leu Glu Thr Thr Leu Thr Arg Ala Arg  
 485 490 495

ggt gga caa ccc cag ttg atc gag gag ata gtc gag gct taa 1530  
 Gly Gly Gln Pro Gln Leu Ile Glu Glu Ile Val Glu Ala  
 500 505

&lt;210&gt; 18

&lt;211&gt; 509

&lt;212&gt; PRT

&lt;213&gt; Brassica napus

&lt;400&gt; 18

Met Val Lys Val Pro His Leu Leu Val Ala Thr Phe Gly Val Leu Leu  
 1 5 10 15

Val Leu Asn Gly Cys Leu Ala Arg Gln Ser Leu Gly Val Pro Pro Gln  
 20 25 30

Leu Gly Asn Ala Cys Asn Leu Asp Asn Leu Asp Val Leu Gln Pro Thr  
 35 40 45

Glu Thr Ile Lys Ser Glu Ala Gly Arg Val Glu Tyr Trp Asp His Asn  
 50 55 60

Asn Pro Gln Ile Arg Cys Ala Gly Val Ser Val Ser Arg Val Ile Ile  
 65 70 75 80

Glu Gln Gly Gly Leu Tyr Leu Pro Thr Phe Phe Ser Ser Pro Lys Ile  
 85 90 95

Ser Tyr Val Val Gln Gly Met Gly Ile Ser Gly Arg Val Val Pro Gly  
 100 105 110

Cys Ala Glu Thr Phe Met Asp Ser Gln Pro Met Gln Gly Gln Gln Gln  
 115 120 125

Gly Gln Pro Trp Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln  
 130 135 140

Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln Gly Gln Gln  
 145 150 155 160  
 Gly Gln Gln Gly Gln Gln Gly Gln Gln Gln Gly Phe Arg Asp Met  
 165 170 175  
 His Gln Lys Val Glu His Val Arg His Gly Asp Ile Ile Ala Ile Thr  
 180 185 190  
 Ala Gly Ser Ser His Trp Ile Tyr Asn Thr Gly Asp Gln Pro Leu Val  
 195 200 205  
 Ile Ile Cys Leu Leu Asp Ile Ala Asn Tyr Gln Asn Gln Leu Asp Arg  
 210 215 220  
 Asn Pro Arg Thr Phe Arg Leu Ala Gly Asn Asn Pro Gln Gly Gly Ser  
 225 230 235 240  
 Gln Gln Gln Gln Gln Gln Gln Gln Asn Met Leu Ser Gly Phe Asp Pro  
 245 250 255  
 Gln Val Leu Ala Gln Ala Leu Lys Ile Asp Val Arg Leu Ala Gln Glu  
 260 265 270  
 Leu Gln Asn Gln Gln Asp Ser Arg Gly Asn Ile Val Arg Val Lys Gly  
 275 280 285  
 Pro Phe Gln Val Val Arg Pro Pro Leu Arg Gln Pro Tyr Glu Ser Glu  
 290 295 300  
 Gln Trp Arg His Pro Arg Gly Pro Pro Gln Ser Pro Gln Asp Asn Gly  
 305 310 315 320  
 Leu Glu Glu Thr Ile Cys Ser Met Arg Thr His Glu Asn Ile Asp Asp  
 325 330 335  
 Pro Ala Arg Ala Asp Val Tyr Lys Pro Asn Leu Gly Arg Val Thr Ser  
 340 345 350  
 Val Asn Ser Tyr Thr Leu Pro Ile Leu Gln Tyr Ile Arg Leu Ser Ala  
 355 360 365  
 Thr Arg Gly Ile Leu Gln Gly Asn Ala Met Val Leu Pro Lys Tyr Asn  
 370 375 380  
 Met Asn Ala Asn Glu Ile Leu Tyr Cys Thr Gln Gly Gln Ala Arg Ile  
 385 390 395 400  
 Gln Val Val Asn Asp Asn Gly Gln Asn Val Leu Asp Gln Gln Val Gln  
 405 410 415  
 Lys Gly Gln Leu Val Val Ile Pro Gln Gly Phe Ala Tyr Val Val Gln  
 420 425 430  
 Ser His Gln Asn Asn Phe Glu Trp Ile Ser Phe Lys Thr Asn Ala Asn  
 435 440 445  
 Ala Met Val Ser Thr Leu Ala Gly Arg Thr Ser Ala Leu Arg Ala Leu  
 450 455 460  
 Pro Leu Glu Val Ile Thr Asn Ala Phe Gln Ile Ser Leu Glu Glu Ala  
 465 470 475 480  
 Arg Arg Ile Lys Phe Asn Thr Leu Glu Thr Thr Leu Thr Arg Ala Arg  
 485 490 495  
 Gly Gly Gln Pro Gln Leu Ile Glu Glu Ile Val Glu Ala  
 500 505

<210> 19  
 <211> 1488  
 <212> DNA  
 <213> Glycine max

<220>  
 <221> CDS  
 <222> (1)..(1485)  
 <223> Gycinin A-1a-B-x subunit

```

<400> 19
atg gcc aag cta gtt ttt tcc ctt tgt ttt ctg ctt ttc agt ggc tgc 48
Met Ala Lys Leu Val Phe Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys
  1          5          10          15

tgc ttc gct ttc agt tcc aga gag cag cct cag caa aac gag tgc cag 96
Cys Phe Ala Phe Ser Ser Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln
          20          25          30

atc caa aaa ctc aat gcc ctc aaa ccg gat aac cgt ata gag tca gaa 144
Ile Gln Lys Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu
          35          40          45

gga ggg ctc att gag aca tgg aac cct aac aac aag cca ttc cag tgt 192
Gly Gly Leu Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys
  50          55          60

gcc ggt gtt gcc ctc tct cgc tgc acc ctc aac cgc aac gcc ctt cgt 240
Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg
  65          70          75          80

aga cct tcc tac acc aac ggt ccc cag gaa atc tac atc caa caa ggt 288
Arg Pro Ser Tyr Thr Asn Gly Pro Gln Glu Ile Tyr Ile Gln Gln Gly
          85          90          95

aag ggt att ttt ggc atg ata tac ccg ggt tgt cct agc aca ttt gaa 336
Lys Gly Ile Phe Gly Met Ile Tyr Pro Gly Cys Pro Ser Thr Phe Glu
          100          105          110

gag cct caa caa cct caa caa aga gga caa agc agc aga cca caa gac 384
Glu Pro Gln Gln Pro Gln Gln Arg Gly Gln Ser Ser Arg Pro Gln Asp
          115          120          125

cgt cac cag aag atc tat aac ttc aga gag ggt gat ttg atc gca gtg 432
Arg His Gln Lys Ile Tyr Asn Phe Arg Glu Gly Asp Leu Ile Ala Val
          130          135          140

cct act ggt gtt gca tgg tgg atg tac aac aat gaa gac act cct gtt 480
Pro Thr Gly Val Ala Trp Trp Met Tyr Asn Asn Glu Asp Thr Pro Val
          145          150          155          160

gtt gcc gtt tct att att gac acc aac agc ttg gag aac cag ctc gac 528
Val Ala Val Ser Ile Ile Asp Thr Asn Ser Leu Glu Asn Gln Leu Asp
          165          170          175

cag atg cct agg aga ttc tat ctt gct ggg aac caa gag caa gag ttt 576
Gln Met Pro Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe
          180          185          190

cta aaa tat cag caa gag caa gga ggt cat caa agc cag aaa gga aag 624
Leu Lys Tyr Gln Gln Glu Gln Gly Gly His Gln Ser Gln Lys Gly Lys
          195          200          205

cat cag caa gaa gaa gaa aac gaa gga ggc agc ata ttg agt ggc ttc 672
His Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe
          210          215          220

```

## 23

acc ctg gaa ttc ttg gaa cat gca ttc agc gtg gac aag cag ata gcg	720
Thr Leu Glu Phe Leu His Ala Phe Ser Val Asp Lys Gln Ile Ala	
225 230 235 240	
aaa aac cta caa gga gag aac gaa ggg gaa gac aag gga gcc att gtg	768
Lys Asn Leu Gln Gly Glu Asn Glu Gly Glu Asp Lys Gly Ala Ile Val	
245 250 255	
aca gtg aaa gga ggt ctg agc gtg ata aaa cca ccc acg gac gag cag	816
Thr Val Lys Gly Gly Leu Ser Val Ile Lys Pro Pro Thr Asp Glu Gln	
260 265 270	
caa caa aga ccc cag gaa gag gaa gaa gaa gaa gag gat gag aag cca	864
Gln Gln Arg Pro Gln Glu Glu Glu Glu Glu Glu Glu Glu Lys Pro	
275 280 285	
cag tgc aag ggt aaa gac aaa cac tgc caa cgc ccc cga gga agc caa	912
Gln Cys Lys Gly Lys Asp Lys His Cys Gln Arg Pro Arg Gly Ser Gln	
290 295 300	
agc aaa agc aga aga aat ggc att gac gag acc ata tgc acc atg aga	960
Ser Lys Ser Arg Arg Asn Gly Ile Asp Glu Thr Ile Cys Thr Met Arg	
305 310 315 320	
ctt cgc cac aac att ggc cag act tca tca cct gac atc tac aac cct	1008
Leu Arg His Asn Ile Gly Gln Thr Ser Ser Pro Asp Ile Tyr Asn Pro	
325 330 335	
caa gcc ggt agc gtc aca acc gcc acc agc ctt gac ttc cca gcc ctc	1056
Gln Ala Gly Ser Val Thr Thr Ala Thr Ser Leu Asp Phe Pro Ala Leu	
340 345 350	
tcg tgg ctc aga ctc agt gct gag ttt gga tct ctc cgc aag aat gca	1104
Ser Trp Leu Arg Leu Ser Ala Glu Phe Gly Ser Leu Arg Lys Asn Ala	
355 360 365	
atg ttc gtg cca cac tac aac ctg aac gcg aac agc ata ata tac gca	1152
Met Phe Val Pro His Tyr Asn Leu Asn Ala Asn Ser Ile Ile Tyr Ala	
370 375 380	
ttg aat gga cgg gca ttg ata caa gtg gtg aat tgc aac ggt gag aga	1200
Leu Asn Gly Arg Ala Leu Ile Gln Val Val Asn Cys Asn Gly Glu Arg	
385 390 395 400	
gtg ttt gat gga gag ctg caa gag gga cgg gtg ctg atc gtg cca caa	1248
Val Phe Asp Gly Glu Leu Gln Glu Gly Arg Val Leu Ile Val Pro Gln	
405 410 415	
aac ttt gtg gtg gct gca aga tca cag agt gac aac ttc gag tat gtg	1296
Asn Phe Val Val Ala Ala Arg Ser Gln Ser Asp Asn Phe Glu Tyr Val	
420 425 430	
tca ttc aag acc aat gat aca ccc atg atc ggc act ctt gca ggg gca	1344
Ser Phe Lys Thr Asn Asp Thr Pro Met Ile Gly Thr Leu Ala Gly Ala	
435 440 445	
aac tca ttg ttg aac gca tta cca gag gaa gtg att cag cac act ttc	1392
Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu Val Ile Gln His Thr Phe	
450 455 460	
aac cta aaa agc cag cag gcc agg cag ata aag aac aac aac cct ttc	1440
Asn Leu Lys Ser Gln Gln Ala Arg Gln Ile Lys Asn Asn Asn Pro Phe	
465 470 475 480	
aag ttc ctg gtt cca cct cag gag tct cag aag aga gct gtg gct tag	1488
Lys Phe Leu Val Pro Pro Gln Glu Ser Gln Lys Arg Ala Val Ala	
485 490 495	

<210> 20  
 <211> 495  
 <212> PRT  
 <213> Glycine max  
 <400> 20  
 Met Ala Lys Leu Val Phe Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys  
 1 5 10 15  
 Cys Phe Ala Phe Ser Ser Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln  
 20 25 30  
 Ile Gln Lys Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu  
 35 40 45  
 Gly Gly Leu Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys  
 50 55 60  
 Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg  
 65 70 75 80  
 Arg Pro Ser Tyr Thr Asn Gly Pro Gln Glu Ile Tyr Ile Gln Gln Gly  
 85 90 95  
 Lys Gly Ile Phe Gly Met Ile Tyr Pro Gly Cys Pro Ser Thr Phe Glu  
 100 105 110  
 Glu Pro Gln Gln Pro Gln Gln Arg Gly Gln Ser Ser Arg Pro Gln Asp  
 115 120 125  
 Arg His Gln Lys Ile Tyr Asn Phe Arg Glu Gly Asp Leu Ile Ala Val  
 130 135 140  
 Pro Thr Gly Val Ala Trp Trp Met Tyr Asn Asn Glu Asp Thr Pro Val  
 145 150 155 160  
 Val Ala Val Ser Ile Ile Asp Thr Asn Ser Leu Glu Asn Gln Leu Asp  
 165 170 175  
 Gln Met Pro Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe  
 180 185 190  
 Leu Lys Tyr Gln Gln Glu Gln Gly Gly His Gln Ser Gln Lys Gly Lys  
 195 200 205  
 His Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe  
 210 215 220  
 Thr Leu Glu Phe Leu Glu His Ala Phe Ser Val Asp Lys Gln Ile Ala  
 225 230 235 240  
 Lys Asn Leu Gln Gly Glu Asn Glu Gly Glu Asp Lys Gly Ala Ile Val  
 245 250 255  
 Thr Val Lys Gly Gly Leu Ser Val Ile Lys Pro Pro Thr Asp Glu Gln  
 260 265 270  
 Gln Gln Arg Pro Gln Glu Glu Glu Glu Glu Glu Asp Glu Lys Pro  
 275 280 285  
 Gln Cys Lys Gly Lys Asp Lys His Cys Gln Arg Pro Arg Gly Ser Gln  
 290 295 300  
 Ser Lys Ser Arg Arg Asn Gly Ile Asp Glu Thr Ile Cys Thr Met Arg  
 305 310 315 320  
 Leu Arg His Asn Ile Gly Gln Thr Ser Ser Pro Asp Ile Tyr Asn Pro  
 325 330 335  
 Gln Ala Gly Ser Val Thr Thr Ala Thr Ser Leu Asp Phe Pro Ala Leu  
 340 345 350

## 25

Ser Trp Leu Arg Leu Ser Ala Glu Phe Gly Ser Leu Arg Lys Asn Ala  
 355 360 365  
 Met Phe Val Pro His Tyr Asn Leu Asn Ala Asn Ser Ile Ile Tyr Ala  
 370 375 380  
 Leu Asn Gly Arg Ala Leu Ile Gln Val Val Asn Cys Asn Gly Glu Arg  
 385 390 395 400  
 Val Phe Asp Gly Glu Leu Gln Glu Gly Arg Val Leu Ile Val Pro Gln  
 405 410 415  
 Asn Phe Val Val Ala Ala Arg Ser Gln Ser Asp Asn Phe Glu Tyr Val  
 420 425 430  
 Ser Phe Lys Thr Asn Asp Thr Pro Met Ile Gly Thr Leu Ala Gly Ala  
 435 440 445  
 Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu Val Ile Gln His Thr Phe  
 450 455 460  
 Asn Leu Lys Ser Gln Gln Ala Arg Gln Ile Lys Asn Asn Asn Pro Phe  
 465 470 475 480  
 Lys Phe Leu Val Pro Pro Gln Glu Ser Gln Lys Arg Ala Val Ala  
 485 490 495

<210> 21  
 <211> 1458  
 <212> DNA  
 <213> Glycine max  
 <220>  
 <221> CDS  
 <222> (1)..(1455)  
 <223> glycinin G2 subunit

<400> 21  
 atg gcc aag ctt gtt ctt tcc ctt tgt ttc ctt ctt ttc agt ggc tgc 48  
 Met Ala Lys Leu Val Leu Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys  
 1 5 10 15  
 ttc gct ctg aga gag cag gca cag caa aat gag tgc cag atc caa aag 96  
 Phe Ala Leu Arg Glu Gln Ala Gln Gln Asn Glu Cys Gln Ile Gln Lys  
 20 25 30  
 ctg aat gcc ctc aaa ccg gat aac cgt ata gag tcg gaa ggt ggg ttc 144  
 Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu Gly Gly Phe  
 35 40 45  
 att gag aca tgg aac cct aac aac aag cca ttc cag tgt gcc ggt gtt 192  
 Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys Ala Gly Val  
 50 55 60  
 gcc ctc tct cgc tgc acc ctt aac cgc aat gcc ctt cgt aga cct tcc 240  
 Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg Arg Pro Ser  
 65 70 75 80  
 tac acc aac ggt ccc cag gaa atc tac ata caa caa ggt aat ggt att 288  
 Tyr Thr Asn Gly Pro Gln Glu Ile Tyr Ile Gln Gln Gly Asn Gly Ile  
 85 90 95  
 ttt ggc atg ata ttc ccg ggt tgt cct agc act tat caa gag ccg caa 336  
 Phe Gly Met Ile Phe Pro Gly Cys Pro Ser Thr Tyr Gln Glu Pro Gln  
 100 105 110



## 26

gaa tct cag caa cga gga cga agc cag agg ccc caa gac cgt cac caa	384
Glu Ser Gln Gln Arg Gly Arg Ser Gln Arg Pro Gln Asp Arg His Gln	
115 120 125	
aag gta cat cgc ttc aga gag ggt gat ttg atc gca gtg cct act ggt	432
Lys Val His Arg Phe Arg Glu Gly Asp Leu Ile Ala Val Pro Thr Gly	
130 135 140	
gtt gca tgg tgg atg tac aac aat gaa gac act cct gtt gtt gcc gtt	480
Val Ala Trp Trp Met Tyr Asn Asn Glu Asp Thr Pro Val Val Ala Val	
145 150 155 160	
tct att att gac acc aac agc ttg gag aac cag ctc gac cag atg cct	528
Ser Ile Ile Asp Thr Asn Ser Leu Glu Asn Gln Leu Asp Gln Met Pro	
165 170 175	
agg aga ttc tat ctt gct ggg aac caa gag caa gag ttt cta aaa tat	576
Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe Leu Lys Tyr	
180 185 190	
cag cag cag cag caa gga ggt tcc caa agc cag aaa gga aag caa caa	624
Gln Gln Gln Gln Gln Gly Gly Ser Gln Ser Gln Lys Gly Lys Gln Gln	
195 200 205	
gaa gaa gaa aac gaa gga agc aac ata ttg agt ggc ttc gcc cct gaa	672
Glu Glu Glu Asn Glu Gly Ser Asn Ile Leu Ser Gly Phe Ala Pro Glu	
210 215 220	
ttc ttg aaa gaa gcg ttc ggc gtg aac atg cag ata gtg aga aac cta	720
Phe Leu Lys Glu Ala Phe Gly Val Asn Met Gln Ile Val Arg Asn Leu	
225 230 235 240	
caa ggt gag aac gaa gag gag gat agt gga gcc att gtg aca gtg aaa	768
Gln Gly Glu Asn Glu Glu Glu Asp Ser Gly Ala Ile Val Thr Val Lys	
245 250 255	
gga ggt cta aga gtc aca gct cca gcc atg agg aag cca cag caa gaa	816
Gly Gly Leu Arg Val Thr Ala Pro Ala Met Arg Lys Pro Gln Gln Glu	
260 265 270	
gaa gat gat gat gat gag gaa gag cag cca cag tgc gtg gag aca gac	864
Glu Asp Asp Asp Asp Glu Glu Glu Gln Pro Gln Cys Val Glu Thr Asp	
275 280 285	
aaa ggt tgc caa cgc caa agc aaa agg agc aga aat ggc att gat gag	912
Lys Gly Cys Gln Arg Gln Ser Lys Arg Ser Arg Asn Gly Ile Asp Glu	
290 295 300	
acc att tgc aca atg aga ctt cgc caa aac att ggt cag aat tca tca	960
Thr Ile Cys Thr Met Arg Leu Arg Gln Asn Ile Gly Gln Asn Ser Ser	
305 310 315 320	
cct gac atc tac aac cct caa gct ggt agc atc aca acc gcc acc agc	1008
Pro Asp Ile Tyr Asn Pro Gln Ala Gly Ser Ile Thr Thr Ala Thr Ser	
325 330 335	
ctt gac ttc cca gcc ctc tgg ctt ctc aaa ctc agt gcc cag tat gga	1056
Leu Asp Phe Pro Ala Leu Trp Leu Leu Lys Leu Ser Ala Gln Tyr Gly	
340 345 350	
tca ctc cgc aag aat gct atg ttc gtg cca cac tac acc ctg aac gcg	1104
Ser Leu Arg Lys Asn Ala Met Phe Val Pro His Tyr Thr Leu Asn Ala	
355 360 365	
aac agc ata ata tac gca ttg aat ggg cgg gca ttg gta caa gtg gtg	1152
Asn Ser Ile Ile Tyr Ala Leu Asn Gly Arg Ala Leu Val Gln Val Val	
370 375 380	

```

aat tgc aat ggt gag aga gtg ttt gat gga gag ctg caa gag gga ggg 1200
Asn Cys Asn Gly Glu Arg Val Phe Asp Gly Glu Leu Gln Glu Gly Gly
385 390 395 400

gtg ctg atc gtt cca caa aac ttt gcg gtg gct gca aaa tcc cag agc 1248
Val Leu Ile Val Pro Gln Asn Phe Ala Val Ala Ala Lys Ser Gln Ser
405 410 415

gat aac ttt gag tat gtg tca ttc aag acc aat gat aga ccc tcg atc 1296
Asp Asn Phe Glu Tyr Val Ser Phe Lys Thr Asn Asp Arg Pro Ser Ile
420 425 430

gga aac ctt gca ggg gca aac tca ttg ttg aac gca ttg cca gag gaa 1344
Gly Asn Leu Ala Gly Ala Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu
435 440 445

gtg att cag cac act ttt aac cta aag agc cag cag gcc agg cag gtg 1392
Val Ile Gln His Thr Phe Asn Leu Lys Ser Gln Gln Ala Arg Gln Val
450 455 460

aag aac aac aac cct ttc agc ttc ctt gtt cca cct cag gag tct cag 1440
Lys Asn Asn Asn Pro Phe Ser Phe Leu Val Pro Pro Gln Glu Ser Gln
465 470 475 480

agg aga gct gtg gct tag 1458
Arg Arg Ala Val Ala
485

```

&lt;210&gt; 22

&lt;211&gt; 485

&lt;212&gt; PRT

&lt;213&gt; Glycine max

&lt;400&gt; 22

```

Met Ala Lys Leu Val Leu Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys
1 5 10 15

Phe Ala Leu Arg Glu Gln Ala Gln Gln Asn Glu Cys Gln Ile Gln Lys
20 25 30

Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu Gly Gly Phe
35 40 45

Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys Ala Gly Val
50 55 60

Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg Arg Pro Ser
65 70 75 80

Tyr Thr Asn Gly Pro Gln Glu Ile Tyr Ile Gln Gln Gly Asn Gly Ile
85 90 95

Phe Gly Met Ile Phe Pro Gly Cys Pro Ser Thr Tyr Gln Glu Pro Gln
100 105 110

Glu Ser Gln Gln Arg Gly Arg Ser Gln Arg Pro Gln Asp Arg His Gln
115 120 125

Lys Val His Arg Phe Arg Glu Gly Asp Leu Ile Ala Val Pro Thr Gly
130 135 140

Val Ala Trp Trp Met Tyr Asn Asn Glu Asp Thr Pro Val Val Ala Val
145 150 155 160

Ser Ile Ile Asp Thr Asn Ser Leu Glu Asn Gln Leu Asp Gln Met Pro
165 170 175

Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe Leu Lys Tyr
180 185 190

```

## 28

Gln Gln Gln Gln Gln Gly Gly Ser Gln Ser Gln Lys Gly Lys Gln Gln  
           195                          200                          205  
 Glu Glu Glu Asn Glu Gly Ser Asn Ile Leu Ser Gly Phe Ala Pro Glu  
           210                          215                          220  
 Phe Leu Lys Glu Ala Phe Gly Val Asn Met Gln Ile Val Arg Asn Leu  
   225                          230                          235                          240  
 Gln Gly Glu Asn Glu Glu Glu Asp Ser Gly Ala Ile Val Thr Val Lys  
                           245                          250                          255  
 Gly Gly Leu Arg Val Thr Ala Pro Ala Met Arg Lys Pro Gln Gln Glu  
                           260                          265                          270  
 Glu Asp Asp Asp Asp Glu Glu Glu Gln Pro Gln Cys Val Glu Thr Asp  
                           275                          280                          285  
 Lys Gly Cys Gln Arg Gln Ser Lys Arg Ser Arg Asn Gly Ile Asp Glu  
           290                          295                          300  
 Thr Ile Cys Thr Met Arg Leu Arg Gln Asn Ile Gly Gln Asn Ser Ser  
   305                          310                          315                          320  
 Pro Asp Ile Tyr Asn Pro Gln Ala Gly Ser Ile Thr Thr Ala Thr Ser  
                           325                          330                          335  
 Leu Asp Phe Pro Ala Leu Trp Leu Leu Lys Leu Ser Ala Gln Tyr Gly  
                           340                          345                          350  
 Ser Leu Arg Lys Asn Ala Met Phe Val Pro His Tyr Thr Leu Asn Ala  
                           355                          360                          365  
 Asn Ser Ile Ile Tyr Ala Leu Asn Gly Arg Ala Leu Val Gln Val Val  
           370                          375                          380  
 Asn Cys Asn Gly Glu Arg Val Phe Asp Gly Glu Leu Gln Glu Gly Gly  
   385                          390                          395                          400  
 Val Leu Ile Val Pro Gln Asn Phe Ala Val Ala Ala Lys Ser Gln Ser  
                           405                          410                          415  
 Asp Asn Phe Glu Tyr Val Ser Phe Lys Thr Asn Asp Arg Pro Ser Ile  
                           420                          425                          430  
 Gly Asn Leu Ala Gly Ala Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu  
           435                          440                          445  
 Val Ile Gln His Thr Phe Asn Leu Lys Ser Gln Gln Ala Arg Gln Val  
           450                          455                          460  
 Lys Asn Asn Asn Pro Phe Ser Phe Leu Val Pro Pro Gln Glu Ser Gln  
   465                          470                          475                          480  
 Arg Arg Ala Val Ala  
                           485

&lt;210&gt; 23

&lt;211&gt; 1689

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1686)

&lt;223&gt; glycinin A5A4B3 subunits

&lt;400&gt; 23

atg ggg aag ccc ttc act ctc tct ctt tct tcc ctt tgc ttg cta ctc 48

Met	Gly	Lys	Pro	Phe	Thr	Leu	Ser	Leu	Ser	Ser	Leu	Cys	Leu	Leu	Leu	
1				5					10					15		
ttg	tcg	agt	gca	tgc	ttt	gct	att	agc	tcc	agc	aag	ctc	aac	gag	tgc	96
Leu	Ser	Ser	Ala	Cys	Phe	Ala	Ile	Ser	Ser	Ser	Lys	Leu	Asn	Glu	Cys	
			20					25					30			
caa	ctc	aac	aac	ctc	aac	gcg	ttg	gaa	ccc	gac	cac	cgc	ggt	gag	tcc	144
Gln	Leu	Asn	Asn	Leu	Asn	Ala	Leu	Glu	Pro	Asp	His	Arg	Val	Glu	Ser	
		35					40				45					
gaa	ggg	ggg	ttg	att	caa	aca	tgg	aac	tct	caa	cac	cct	gag	ctg	aaa	192
Glu	Gly	Gly	Leu	Ile	Gln	Thr	Trp	Asn	Ser	Gln	His	Pro	Glu	Leu	Lys	
	50					55				60						
tgc	gcc	ggg	gtc	act	gtt	tcc	aaa	ctc	acc	ctc	aac	cgc	aat	ggc	ctc	240
Cys	Ala	Gly	Val	Thr	Val	Ser	Lys	Leu	Thr	Leu	Asn	Arg	Asn	Gly	Leu	
	65				70					75					80	
cac	tcg	cca	tct	tac	tca	cct	tat	ccc	cgg	atg	atc	atc	atc	gcc	caa	288
His	Ser	Pro	Ser	Tyr	Ser	Pro	Tyr	Pro	Arg	Met	Ile	Ile	Ile	Ala	Gln	
				85				90						95		
ggg	aaa	gga	gca	ctt	gga	gtt	gca	att	cca	gga	tgt	cct	gag	acg	ttt	336
Gly	Lys	Gly	Ala	Leu	Gly	Val	Ala	Ile	Pro	Gly	Cys	Pro	Glu	Thr	Phe	
			100					105					110			
gag	gag	cca	caa	gaa	caa	tca	aac	aga	aga	ggc	tca	agg	tcg	cag	aag	384
Glu	Glu	Pro	Gln	Glu	Gln	Ser	Asn	Arg	Arg	Gly	Ser	Arg	Ser	Gln	Lys	
		115					120					125				
cag	cag	cta	cag	gac	agt	cac	cag	aag	att	cgt	cac	ttc	aat	gaa	gga	432
Gln	Gln	Leu	Gln	Asp	Ser	His	Gln	Lys	Ile	Arg	His	Phe	Asn	Glu	Gly	
	130					135				140						
gac	gta	ctc	gtg	att	cct	cct	agt	gtt	cct	tac	tgg	acc	tat	aac	act	480
Asp	Val	Leu	Val	Ile	Pro	Pro	Ser	Val	Pro	Tyr	Trp	Thr	Tyr	Asn	Thr	
	145				150					155					160	
ggc	gat	gaa	cca	gtt	gtt	gcc	atc	agt	ctt	ctt	gac	acc	tct	aac	ttc	528
Gly	Asp	Glu	Pro	Val	Val	Ala	Ile	Ser	Leu	Leu	Asp	Thr	Ser	Asn	Phe	
				165					170					175		
aat	aac	cag	ctt	gat	caa	acc	cct	agg	gta	ttt	tac	ctt	gct	ggg	aac	576
Asn	Asn	Gln	Leu	Asp	Gln	Thr	Pro	Arg	Val	Phe	Tyr	Leu	Ala	Gly	Asn	
			180					185					190			
cca	gat	ata	gag	tac	cca	gag	acc	atg	caa	caa	caa	caa	cag	cag	aaa	624
Pro	Asp	Ile	Glu	Tyr	Pro	Glu	Thr	Met	Gln	Gln	Gln	Gln	Gln	Gln	Lys	
		195					200					205				
agt	cat	ggg	gga	cgc	aag	cag	ggg	caa	cac	cag	cag	gag	gaa	gag	gaa	672
Ser	His	Gly	Gly	Arg	Lys	Gln	Gly	Gln	His	Gln	Gln	Glu	Glu	Glu	Glu	
	210					215					220					
gaa	ggg	ggc	agc	gtg	ctc	agt	ggc	ttc	agc	aaa	cac	ttc	ttg	gca	caa	720
Glu	Gly	Gly	Ser	Val	Leu	Ser	Gly	Phe	Ser	Lys	His	Phe	Leu	Ala	Gln	
	225				230					235				240		
tcc	ttc	aac	acc	aac	gag	gac	ata	gct	gag	aaa	ctt	gag	tct	cca	gac	768
Ser	Phe	Asn	Thr	Asn	Glu	Asp	Ile	Ala	Glu	Lys	Leu	Glu	Ser	Pro	Asp	
				245					250					255		
gac	gaa	agg	aag	cag	atc	gtg	aca	gtg	gaa	gga	ggg	ctc	agc	gtt	atc	816
Asp	Glu	Arg	Lys	Gln	Ile	Val	Thr	Val	Glu	Gly	Gly	Leu	Ser	Val	Ile	
			260					265					270			

## 30

agc	ccc	aag	tgg	caa	gaa	caa	caa	gat	gaa	gat	gaa	gat	gaa	gac	gaa	864
Ser	Pro	Lys	Trp	Gln	Glu	Gln	Gln	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Glu	
		275					280					285				
gat	gat	gaa	gat	gaa	caa	att	ccc	tct	cac	cct	cct	cgc	cga	cca	agc	912
Asp	Asp	Glu	Asp	Glu	Gln	Ile	Pro	Ser	His	Pro	Pro	Arg	Arg	Pro	Ser	
		290				295					300					
cat	gga	aag	cgt	gaa	caa	gac	gag	gac	gag	gac	gaa	gat	gaa	gat	aaa	960
His	Gly	Lys	Arg	Glu	Gln	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Lys	
					310					315					320	
cct	cgt	cct	agt	cga	cca	agc	caa	gga	aag	cgg	aac	aag	aca	gga	cag	1008
Pro	Arg	Pro	Ser	Arg	Pro	Ser	Gln	Gly	Lys	Arg	Asn	Lys	Thr	Gly	Gln	
				325					330					335		
gac	gag	gac	gaa	gat	gaa	gat	gaa	gat	caa	cct	cgc	aag	agc	cgc	gaa	1056
Asp	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Gln	Pro	Arg	Lys	Ser	Arg	Glu	
			340					345					350			
tgg	aga	tcg	aaa	aag	aca	caa	ccc	aga	aga	cct	aga	caa	gaa	gaa	cca	1104
Trp	Arg	Ser	Lys	Lys	Thr	Gln	Pro	Arg	Arg	Pro	Arg	Gln	Glu	Glu	Pro	
		355					360					365				
cgt	gaa	aga	gga	tgc	gag	aca	aga	aac	ggg	gtt	gag	gaa	aat	atc	tgc	1152
Arg	Glu	Arg	Gly	Cys	Glu	Thr	Arg	Asn	Gly	Val	Glu	Glu	Asn	Ile	Cys	
		370				375					380					
acc	ttg	aag	ctt	cac	gag	aac	att	gct	cgc	cct	tca	cgc	gct	gac	ttc	1200
Thr	Leu	Lys	Leu	His	Glu	Asn	Ile	Ala	Arg	Pro	Ser	Arg	Ala	Asp	Phe	
					390					395					400	
tac	aac	cct	aaa	gct	ggt	cgc	att	agt	acc	ctc	aac	agc	ctc	acc	ctc	1248
Tyr	Asn	Pro	Lys	Ala	Gly	Arg	Ile	Ser	Thr	Leu	Asn	Ser	Leu	Thr	Leu	
				405					410					415		
cca	gcc	ctc	cgc	caa	ttc	caa	ctc	agt	gcc	caa	tat	gtt	gtc	ctc	tac	1296
Pro	Ala	Leu	Arg	Gln	Phe	Gln	Leu	Ser	Ala	Gln	Tyr	Val	Val	Leu	Tyr	
			420					425					430			
aag	aat	gga	att	tac	tct	cca	cat	tgg	aat	ctg	aat	gca	aac	agt	gtg	1344
Lys	Asn	Gly	Ile	Tyr	Ser	Pro	His	Trp	Asn	Leu	Asn	Ala	Asn	Ser	Val	
		435					440					445				
atc	tat	gtg	act	cga	gga	caa	gga	aag	gtt	aga	gtt	gtg	aac	tgc	caa	1392
Ile	Tyr	Val	Thr	Arg	Gly	Gln	Gly	Lys	Val	Arg	Val	Val	Asn	Cys	Gln	
		450				455					460					
ggg	aat	gca	gtg	ttc	gac	ggt	gag	ctt	agg	agg	gga	caa	ttg	ctg	gtg	1440
Gly	Asn	Ala	Val	Phe	Asp	Gly	Glu	Leu	Arg	Arg	Gly	Gln	Leu	Leu	Val	
					470				475					480		
gta	cca	cag	aac	ttc	gtg	gtg	gcg	gag	caa	gcc	gga	gaa	caa	gga	ttc	1488
Val	Pro	Gln	Asn	Phe	Val	Val	Ala	Glu	Gln	Ala	Gly	Glu	Gln	Gly	Phe	
				485				490						495		
gaa	tac	ata	gta	ttc	aag	aca	cac	cac	aac	gca	gtc	act	agc	tac	ttg	1536
Glu	Tyr	Ile	Val	Phe	Lys	Thr	His	His	Asn	Ala	Val	Thr	Ser	Tyr	Leu	
			500					505					510			
aag	gat	gtg	ttt	agg	gca	att	ccc	tca	gag	gtt	ctt	gcc	cat	tct	tac	1584
Lys	Asp	Val	Phe	Arg	Ala	Ile	Pro	Ser	Glu	Val	Leu	Ala	His	Ser	Tyr	
		515					520					525				
aac	ctt	cga	cag	agt	caa	gtg	tct	gag	ctt	aag	tat	gaa	gga	aat	tgg	1632
Asn	Leu	Arg	Gln	Ser	Gln	Val	Ser	Glu	Leu	Lys	Tyr	Glu	Gly	Asn	Trp	
		530				535					540					

## 31

ggt cct ttg gtc aac cct gag tct caa caa ggc tca ccc cgt gtt aaa 1680  
 Gly Pro Leu Val Asn Pro Glu Ser Gln Gln Gly Ser Pro Arg Val Lys  
 545 550 555 560

gtc gca taa 1689  
 Val Ala

<210> 24

<211> 562

<212> PRT

<213> Glycine max

<400> 24

Met Gly Lys Pro Phe Thr Leu Ser Leu Ser Ser Leu Cys Leu Leu Leu  
 1 5 10 15

Leu Ser Ser Ala Cys Phe Ala Ile Ser Ser Ser Lys Leu Asn Glu Cys  
 20 25 30

Gln Leu Asn Asn Leu Asn Ala Leu Glu Pro Asp His Arg Val Glu Ser  
 35 40 45

Glu Gly Gly Leu Ile Gln Thr Trp Asn Ser Gln His Pro Glu Leu Lys  
 50 55 60

Cys Ala Gly Val Thr Val Ser Lys Leu Thr Leu Asn Arg Asn Gly Leu  
 65 70 75 80

His Ser Pro Ser Tyr Ser Pro Tyr Pro Arg Met Ile Ile Ile Ala Gln  
 85 90 95

Gly Lys Gly Ala Leu Gly Val Ala Ile Pro Gly Cys Pro Glu Thr Phe  
 100 105 110

Glu Glu Pro Gln Glu Gln Ser Asn Arg Arg Gly Ser Arg Ser Gln Lys  
 115 120 125

Gln Gln Leu Gln Asp Ser His Gln Lys Ile Arg His Phe Asn Glu Gly  
 130 135 140

Asp Val Leu Val Ile Pro Pro Ser Val Pro Tyr Trp Thr Tyr Asn Thr  
 145 150 155 160

Gly Asp Glu Pro Val Val Ala Ile Ser Leu Leu Asp Thr Ser Asn Phe  
 165 170 175

Asn Asn Gln Leu Asp Gln Thr Pro Arg Val Phe Tyr Leu Ala Gly Asn  
 180 185 190

Pro Asp Ile Glu Tyr Pro Glu Thr Met Gln Gln Gln Gln Gln Lys  
 195 200 205

Ser His Gly Gly Arg Lys Gln Gly Gln His Gln Gln Glu Glu Glu Glu  
 210 215 220

Glu Gly Gly Ser Val Leu Ser Gly Phe Ser Lys His Phe Leu Ala Gln  
 225 230 235 240

Ser Phe Asn Thr Asn Glu Asp Ile Ala Glu Lys Leu Glu Ser Pro Asp  
 245 250 255

Asp Glu Arg Lys Gln Ile Val Thr Val Glu Gly Gly Leu Ser Val Ile  
 260 265 270

Ser Pro Lys Trp Gln Glu Gln Gln Asp Glu Asp Glu Asp Glu Asp Glu  
 275 280 285

Asp Asp Glu Asp Glu Gln Ile Pro Ser His Pro Pro Arg Arg Pro Ser  
 290 295 300

32

His Gly Lys Arg Glu Gln Asp Glu Asp Glu Asp Glu Asp Glu Asp Lys  
 305 310 315 320  
 Pro Arg Pro Ser Arg Pro Ser Gln Gly Lys Arg Asn Lys Thr Gly Gln  
 325 330 335  
 Asp Glu Asp Glu Asp Glu Asp Glu Asp Gln Pro Arg Lys Ser Arg Glu  
 340 345 350  
 Trp Arg Ser Lys Lys Thr Gln Pro Arg Arg Pro Arg Gln Glu Glu Pro  
 355 360 365  
 Arg Glu Arg Gly Cys Glu Thr Arg Asn Gly Val Glu Glu Asn Ile Cys  
 370 375 380  
 Thr Leu Lys Leu His Glu Asn Ile Ala Arg Pro Ser Arg Ala Asp Phe  
 385 390 395 400  
 Tyr Asn Pro Lys Ala Gly Arg Ile Ser Thr Leu Asn Ser Leu Thr Leu  
 405 410 415  
 Pro Ala Leu Arg Gln Phe Gln Leu Ser Ala Gln Tyr Val Val Leu Tyr  
 420 425 430  
 Lys Asn Gly Ile Tyr Ser Pro His Trp Asn Leu Asn Ala Asn Ser Val  
 435 440 445  
 Ile Tyr Val Thr Arg Gly Gln Gly Lys Val Arg Val Val Asn Cys Gln  
 450 455 460  
 Gly Asn Ala Val Phe Asp Gly Glu Leu Arg Arg Gly Gln Leu Leu Val  
 465 470 475 480  
 Val Pro Gln Asn Phe Val Val Ala Glu Gln Ala Gly Glu Gln Gly Phe  
 485 490 495  
 Glu Tyr Ile Val Phe Lys Thr His His Asn Ala Val Thr Ser Tyr Leu  
 500 505 510  
 Lys Asp Val Phe Arg Ala Ile Pro Ser Glu Val Leu Ala His Ser Tyr  
 515 520 525  
 Asn Leu Arg Gln Ser Gln Val Ser Glu Leu Lys Tyr Glu Gly Asn Trp  
 530 535 540  
 Gly Pro Leu Val Asn Pro Glu Ser Gln Gln Gly Ser Pro Arg Val Lys  
 545 550 555 560  
 Val Ala

&lt;210&gt; 25

&lt;211&gt; 1551

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1548)

&lt;223&gt; glycinin A3-B4 subunit

&lt;400&gt; 25

atg ggg aag ccc ttc ttc act ctc tct ctt tct tcc ctt tgc ttg cta 48  
 Met Gly Lys Pro Phe Phe Thr Leu Ser Leu Ser Ser Leu Cys Leu Leu  
 1 5 10 15

ctc ttg tcg agt gca tgc ttt gct att acc tcc agc aag ttc aac gag 96  
 Leu Leu Ser Ser Ala Cys Phe Ala Ile Thr Ser Ser Lys Phe Asn Glu  
 20 25 30

tgc	caa	ctc	aac	aac	ctc	aac	gcg	ttg	gaa	ccc	gac	cac	cgc	gtt	gag	144
Cys	Gln	Leu	Asn	Asn	Leu	Asn	Ala	Leu	Glu	Pro	Asp	His	Arg	Val	Glu	
	35						40					45				
tcc	gaa	ggt	ggt	ctt	att	gaa	aca	tgg	aac	tct	caa	cac	cct	gag	ctg	192
Ser	Glu	Gly	Gly	Leu	Ile	Glu	Thr	Trp	Asn	Ser	Gln	His	Pro	Glu	Leu	
	50					55					60					
caa	tgc	gcc	ggt	gtc	act	gtt	tcc	aaa	cgc	acc	ctc	aac	cgc	aac	ggc	240
Gln	Cys	Ala	Gly	Val	Thr	Val	Ser	Lys	Arg	Thr	Leu	Asn	Arg	Asn	Gly	
	65				70					75					80	
tcc	cac	ttg	cca	tct	tac	tta	cct	tat	ccc	caa	atg	atc	att	gtc	gtt	288
Ser	His	Leu	Pro	Ser	Tyr	Leu	Pro	Tyr	Pro	Gln	Met	Ile	Ile	Val	Val	
				85					90					95		
caa	ggg	aag	gga	gca	att	gga	ttt	gca	ttt	ccg	gga	tgt	ccc	gag	acg	336
Gln	Gly	Lys	Gly	Ala	Ile	Gly	Phe	Ala	Phe	Pro	Gly	Cys	Pro	Glu	Thr	
		100					105						110			
ttt	gag	aag	cca	caa	caa	caa	tca	agc	aga	aga	ggc	tca	agg	tca	cag	384
Phe	Glu	Lys	Pro	Gln	Gln	Gln	Ser	Ser	Arg	Arg	Gly	Ser	Arg	Ser	Gln	
		115				120						125				
cag	caa	cta	caa	gac	agt	cac	cag	aag	att	cgt	cac	ttc	aat	gaa	gga	432
Gln	Gln	Leu	Gln	Asp	Ser	His	Gln	Lys	Ile	Arg	His	Phe	Asn	Glu	Gly	
	130					135					140					
gac	gta	cta	gtg	att	cct	ctt	ggt	gtt	cct	tac	tgg	acc	tat	aac	act	480
Asp	Val	Leu	Val	Ile	Pro	Leu	Gly	Val	Pro	Tyr	Trp	Thr	Tyr	Asn	Thr	
	145				150					155					160	
ggc	gat	gaa	cca	gtt	gtt	gcc	atc	agt	cct	ctt	gac	acc	tcc	aac	ttc	528
Gly	Asp	Glu	Pro	Val	Val	Ala	Ile	Ser	Pro	Leu	Asp	Thr	Ser	Asn	Phe	
				165					170					175		
aac	aat	cag	ctt	gat	caa	aac	ccc	aga	gta	ttt	tac	ctt	gct	ggg	aac	576
Asn	Asn	Gln	Leu	Asp	Gln	Asn	Pro	Arg	Val	Phe	Tyr	Leu	Ala	Gly	Asn	
		180						185					190			
cca	gat	ata	gag	cat	ccc	gag	acc	atg	caa	caa	cag	cag	cag	cag	aag	624
Pro	Asp	Ile	Glu	His	Pro	Glu	Thr	Met	Gln	Gln	Gln	Gln	Gln	Gln	Lys	
		195					200					205				
agt	cat	ggt	gga	cgc	aag	cag	ggg	caa	cac	cga	cag	cag	cag	gag	gaa	672
Ser	His	Gly	Gly	Arg	Lys	Gln	Gly	Gln	His	Arg	Gln	Gln	Glu	Glu	Glu	
	210					215					220					
ggt	ggc	agt	gtg	ctc	agt	ggc	ttc	agc	aaa	cat	ttc	tta	gca	caa	tcc	720
Gly	Gly	Ser	Val	Leu	Ser	Gly	Phe	Ser	Lys	His	Phe	Leu	Ala	Gln	Ser	
	225				230					235					240	
ttc	aac	acc	aac	gag	gac	aca	gct	gag	aaa	ctt	cgg	tct	cca	gat	gac	768
Phe	Asn	Thr	Asn	Glu	Asp	Thr	Ala	Glu	Lys	Leu	Arg	Ser	Pro	Asp	Asp	
			245						250					255		
gaa	agg	aag	cag	atc	gtg	aca	gtg	gag	gga	ggc	ctc	agc	gtt	atc	agc	816
Glu	Arg	Lys	Gln	Ile	Val	Thr	Val	Glu	Gly	Gly	Leu	Ser	Val	Ile	Ser	
		260						265					270			
ccc	aag	tgg	caa	gaa	caa	gaa	gac	gaa	gac	gaa	gac	gaa	gac	gaa	gaa	864
Pro	Lys	Trp	Gln	Glu	Gln	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Glu	
		275					280						285			
tat	gga	cgg	acg	ccc	tct	tat	cct	cca	cga	cga	cca	agc	cat	gga	aag	912
Tyr	Gly	Arg	Thr	Pro	Ser	Tyr	Pro	Pro	Arg	Arg	Pro	Ser	His	Gly	Lys	
	290					295					300					



## 34

cat	gaa	gat	gac	gag	gac	gag	gac	gaa	gaa	gaa	gat	caa	cct	cgt	cct	960
His	Glu	Asp	Asp	Glu	Asp	Glu	Asp	Glu	Glu	Glu	Asp	Gln	Pro	Arg	Pro	
305						310				315					320	
gat	cac	cct	cca	cag	cga	cca	agc	agg	ccc	gaa	caa	caa	gaa	cca	cgt	1008
Asp	His	Pro	Pro	Gln	Arg	Pro	Ser	Arg	Pro	Glu	Gln	Gln	Glu	Pro	Arg	
				325					330						335	
gga	aga	gga	tgt	cag	act	aga	aat	ggg	gtt	gag	gaa	aat	att	tgc	acc	1056
Gly	Arg	Gly	Cys	Gln	Thr	Arg	Asn	Gly	Val	Glu	Glu	Asn	Ile	Cys	Thr	
			340					345					350			
atg	aag	ctt	cac	gag	aac	att	gct	cgc	cct	tca	cgt	gct	gac	ttc	tac	1104
Met	Lys	Leu	His	Glu	Asn	Ile	Ala	Arg	Pro	Ser	Arg	Ala	Asp	Phe	Tyr	
		355					360					365				
aac	cca	aaa	gct	ggt	cgc	att	agc	acc	ctc	aac	agt	ctc	acc	ctc	cca	1152
Asn	Pro	Lys	Ala	Gly	Arg	Ile	Ser	Thr	Leu	Asn	Ser	Leu	Thr	Leu	Pro	
	370					375					380					
gcc	ctc	cgc	caa	ttc	gga	ctc	agt	gcc	caa	tat	gtt	gtc	ctc	tac	agg	1200
Ala	Leu	Arg	Gln	Phe	Gly	Leu	Ser	Ala	Gln	Tyr	Val	Val	Leu	Tyr	Arg	
385					390				395						400	
aat	gga	att	tac	tct	cca	gat	tgg	aac	ttg	aac	gcg	aac	agt	gtg	acg	1248
Asn	Gly	Ile	Tyr	Ser	Pro	Asp	Trp	Asn	Leu	Asn	Ala	Asn	Ser	Val	Thr	
				405				410						415		
atg	act	cga	ggg	aaa	gga	aga	gtt	aga	gtg	gtg	aac	tgc	caa	ggg	aat	1296
Met	Thr	Arg	Gly	Lys	Gly	Arg	Val	Arg	Val	Val	Asn	Cys	Gln	Gly	Asn	
			420					425					430			
gca	gtg	ttc	gac	ggt	gag	cta	agg	agg	gga	caa	ttg	cta	gtg	gtg	ccg	1344
Ala	Val	Phe	Asp	Gly	Glu	Leu	Arg	Arg	Gly	Gln	Leu	Leu	Val	Val	Pro	
		435					440					445				
cag	aac	ccc	gcg	gtg	gct	gag	caa	ggg	gga	gaa	caa	gga	ttg	gaa	tat	1392
Gln	Asn	Pro	Ala	Val	Ala	Glu	Gln	Gly	Gly	Glu	Gln	Gly	Leu	Glu	Tyr	
	450					455					460					
gta	gtg	ttc	aag	aca	cac	cac	aac	gcc	gtg	agc	agc	tac	att	aag	gat	1440
Val	Val	Phe	Lys	Thr	His	His	Asn	Ala	Val	Ser	Ser	Tyr	Ile	Lys	Asp	
465					470				475						480	
gtg	ttt	agg	gta	atc	cct	tcg	gag	gtt	ctt	tcc	aat	tct	tac	aac	ctt	1488
Val	Phe	Arg	Val	Ile	Pro	Ser	Glu	Val	Leu	Ser	Asn	Ser	Tyr	Asn	Leu	
				485				490						495		
ggc	cag	agt	caa	gtg	cgt	cag	ctc	aag	tat	caa	gga	aac	tcc	ggc	cct	1536
Gly	Gln	Ser	Gln	Val	Arg	Gln	Leu	Lys	Tyr	Gln	Gly	Asn	Ser	Gly	Pro	
			500					505					510			
ttg	gtc	aac	cca	taa												1551
Leu	Val	Asn	Pro													
			515													
<210>	26															
<211>	516															
<212>	PRT															
<213>	Glycine max															
<400>	26															
Met	Gly	Lys	Pro	Phe	Phe	Thr	Leu	Ser	Leu	Ser	Ser	Leu	Cys	Leu	Leu	
1				5					10				15			
Leu	Leu	Ser	Ser	Ala	Cys	Phe	Ala	Ile	Thr	Ser	Ser	Lys	Phe	Asn	Glu	
			20					25					30			

## 35

Cys	Gln	Leu	Asn	Asn	Leu	Asn	Ala	Leu	Glu	Pro	Asp	His	Arg	Val	Glu		
		35					40					45					
Ser	Glu	Gly	Gly	Leu	Ile	Glu	Thr	Trp	Asn	Ser	Gln	His	Pro	Glu	Leu		
	50					55					60						
Gln	Cys	Ala	Gly	Val	Thr	Val	Ser	Lys	Arg	Thr	Leu	Asn	Arg	Asn	Gly		
	65				70					75					80		
Ser	His	Leu	Pro	Ser	Tyr	Leu	Pro	Tyr	Pro	Gln	Met	Ile	Ile	Val	Val		
				85					90					95			
Gln	Gly	Lys	Gly	Ala	Ile	Gly	Phe	Ala	Phe	Pro	Gly	Cys	Pro	Glu	Thr		
			100					105					110				
Phe	Glu	Lys	Pro	Gln	Gln	Gln	Ser	Ser	Arg	Arg	Gly	Ser	Arg	Ser	Gln		
	115						120					125					
Gln	Gln	Leu	Gln	Asp	Ser	His	Gln	Lys	Ile	Arg	His	Phe	Asn	Glu	Gly		
	130					135					140						
Asp	Val	Leu	Val	Ile	Pro	Leu	Gly	Val	Pro	Tyr	Trp	Thr	Tyr	Asn	Thr		
145					150					155					160		
Gly	Asp	Glu	Pro	Val	Val	Ala	Ile	Ser	Pro	Leu	Asp	Thr	Ser	Asn	Phe		
				165					170					175			
Asn	Asn	Gln	Leu	Asp	Gln	Asn	Pro	Arg	Val	Phe	Tyr	Leu	Ala	Gly	Asn		
			180					185					190				
Pro	Asp	Ile	Glu	His	Pro	Glu	Thr	Met	Gln	Gln	Gln	Gln	Gln	Gln	Lys		
	195						200						205				
Ser	His	Gly	Gly	Arg	Lys	Gln	Gly	Gln	His	Arg	Gln	Gln	Glu	Glu	Glu		
	210					215					220						
Gly	Gly	Ser	Val	Leu	Ser	Gly	Phe	Ser	Lys	His	Phe	Leu	Ala	Gln	Ser		
225					230					235					240		
Phe	Asn	Thr	Asn	Glu	Asp	Thr	Ala	Glu	Lys	Leu	Arg	Ser	Pro	Asp	Asp		
				245					250					255			
Glu	Arg	Lys	Gln	Ile	Val	Thr	Val	Glu	Gly	Gly	Leu	Ser	Val	Ile	Ser		
			260					265						270			
Pro	Lys	Trp	Gln	Glu	Gln	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Asp	Glu	Glu		
	275						280					285					
Tyr	Gly	Arg	Thr	Pro	Ser	Tyr	Pro	Pro	Arg	Arg	Pro	Ser	His	Gly	Lys		
	290					295					300						
His	Glu	Asp	Asp	Glu	Asp	Glu	Asp	Glu	Glu	Glu	Asp	Gln	Pro	Arg	Pro		
305					310					315					320		
Asp	His	Pro	Pro	Gln	Arg	Pro	Ser	Arg	Pro	Glu	Gln	Gln	Glu	Pro	Arg		
				325					330					335			
Gly	Arg	Gly	Cys	Gln	Thr	Arg	Asn	Gly	Val	Glu	Glu	Asn	Ile	Cys	Thr		
			340					345					350				
Met	Lys	Leu	His	Glu	Asn	Ile	Ala	Arg	Pro	Ser	Arg	Ala	Asp	Phe	Tyr		
	355						360					365					
Asn	Pro	Lys	Ala	Gly	Arg	Ile	Ser	Thr	Leu	Asn	Ser	Leu	Thr	Leu	Pro		
	370					375					380						
Ala	Leu	Arg	Gln	Phe	Gly	Leu	Ser	Ala	Gln	Tyr	Val	Val	Leu	Tyr	Arg		
385					390					395					400		
Asn	Gly	Ile	Tyr	Ser	Pro	Asp	Trp	Asn	Leu	Asn	Ala	Asn	Ser	Val	Thr		
				405					410					415			

Met Thr Arg Gly Lys Gly Arg Val Arg Val Val Asn Cys Gln Gly Asn  
 420 425 430  
 Ala Val Phe Asp Gly Glu Leu Arg Arg Gly Gln Leu Leu Val Val Pro  
 435 440 445  
 Gln Asn Pro Ala Val Ala Glu Gln Gly Gly Glu Gln Gly Leu Glu Tyr  
 450 455 460  
 Val Val Phe Lys Thr His His Asn Ala Val Ser Ser Tyr Ile Lys Asp  
 465 470 475 480  
 Val Phe Arg Val Ile Pro Ser Glu Val Leu Ser Asn Ser Tyr Asn Leu  
 485 490 495  
 Gly Gln Ser Gln Val Arg Gln Leu Lys Tyr Gln Gly Asn Ser Gly Pro  
 500 505 510  
 Leu Val Asn Pro  
 515

<210> 27  
 <211> 1446  
 <212> DNA  
 <213> Glycine max  
 <220>  
 <221> CDS  
 <222> (1)..(1443)  
 <223> glycinin G3 subunit

<400> 27  
 atg gct aag ctt gtt ctt tcc ctt tgt ttt ctg ctt ttc agt ggc tgc 48  
 Met Ala Lys Leu Val Leu Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys  
 1 5 10 15  
 tgc ttc gct ttc agt ttc aga gag cag cca cag caa aac gag tgc cag 96  
 Cys Phe Ala Phe Ser Phe Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln  
 20 25 30  
 atc caa cgc ctc aat gcc cta aaa ccg gat aac cgt ata gag tca gaa 144  
 Ile Gln Arg Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu  
 35 40 45  
 ggt ggc ttc att gag aca tgg aac cct aac aac aag cca ttc cag tgt 192  
 Gly Gly Phe Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys  
 50 55 60  
 gcc ggt gtt gcc ctc tct cgc tgc acc ctc aac cgc aac gcc ctt cgc 240  
 Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg  
 65 70 75 80  
 aga cct tcc tac acc aac gct ccc cag gag atc tac atc caa caa ggt 288  
 Arg Pro Ser Tyr Thr Asn Ala Pro Gln Glu Ile Tyr Ile Gln Gln Gly  
 85 90 95  
 agt ggt att ttt ggc atg ata ttc ccg ggt tgt cct agc aca ttt gaa 336  
 Ser Gly Ile Phe Gly Met Ile Phe Pro Gly Cys Pro Ser Thr Phe Glu  
 100 105 110  
 gag cct caa caa aaa gga caa agc agc agg ccc caa gac cgt cac cag 384  
 Glu Pro Gln Gln Lys Gly Gln Ser Ser Arg Pro Gln Asp Arg His Gln  
 115 120 125  
 aag atc tat cac ttc aga gag ggt gat ttg att gca gtg cca acc ggt 432  
 Lys Ile Tyr His Phe Arg Glu Gly Asp Leu Ile Ala Val Pro Thr Gly  
 130 135 140

ttt gca tac tgg atg tac aac aat gaa gac act cct gtt gtt gcc gtt	480
Phe Ala Tyr Trp Met Tyr Asn Asn Glu Asp Thr Pro Val Val Ala Val	
145 150 155 160	
tct ctt att gac acc aac agc ttc cag aac cag ctc gac cag atg cct	528
Ser Leu Ile Asp Thr Asn Ser Phe Gln Asn Gln Leu Asp Gln Met Pro	
165 170 175	
agg aga ttc tat ctt gct ggg aac caa gag caa gag ttt cta cag tat	576
Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe Leu Gln Tyr	
180 185 190	
cag cca cag aag cag caa gga ggt act caa agc cag aaa gga aag cgt	624
Gln Pro Gln Lys Gln Gln Gly Gly Thr Gln Ser Gln Lys Gly Lys Arg	
195 200 205	
cag caa gaa gaa gaa aac gaa gga ggc agc ata ttg agt ggc ttc gcc	672
Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe Ala	
210 215 220	
ccg gaa ttc ttg gaa cat gcg ttc gtc gtg gac agg cag ata gtg aga	720
Pro Glu Phe Leu Glu His Ala Phe Val Val Asp Arg Gln Ile Val Arg	
225 230 235 240	
aag cta caa ggt gag aac gaa gag gaa gag aag ggt gcc att gtg aca	768
Lys Leu Gln Gly Glu Asn Glu Glu Glu Lys Gly Ala Ile Val Thr	
245 250 255	
gtg aaa gga ggt ctc agc gtg ata agc cca ccc acg gaa gag cag caa	816
Val Lys Gly Gly Leu Ser Val Ile Ser Pro Pro Thr Glu Glu Gln Gln	
260 265 270	
caa aga ccc gag gaa gag gag aag cca gat tgt gac gag aaa gac aaa	864
Gln Arg Pro Glu Glu Glu Glu Lys Pro Asp Cys Asp Glu Lys Asp Lys	
275 280 285	
cat tgc caa agc caa agc aga aat ggc att gac gag acc att tgc aca	912
His Cys Gln Ser Gln Ser Arg Asn Gly Ile Asp Glu Thr Ile Cys Thr	
290 295 300	
atg aga ctt cgc cac aac att ggc cag act tca tca cct gac atc ttc	960
Met Arg Leu Arg His Asn Ile Gly Gln Thr Ser Ser Pro Asp Ile Phe	
305 310 315 320	
aac cct caa gct ggt agc atc aca acc gct acc agc ctc gac ttc cca	1008
Asn Pro Gln Ala Gly Ser Ile Thr Thr Ala Thr Ser Leu Asp Phe Pro	
325 330 335	
gcc ctc tcg tgg ctc aaa ctc agt gcc cag ttt gga tca ctc cgc aag	1056
Ala Leu Ser Trp Leu Lys Leu Ser Ala Gln Phe Gly Ser Leu Arg Lys	
340 345 350	
aat gct atg ttc gtg cca cac tac aac ctg aac gca aac agc ata ata	1104
Asn Ala Met Phe Val Pro His Tyr Asn Leu Asn Ala Asn Ser Ile Ile	
355 360 365	
tac gca ttg aat gga cgg gca ttg gta caa gtg gtg aat tgc aat ggt	1152
Tyr Ala Leu Asn Gly Arg Ala Leu Val Gln Val Asn Cys Asn Gly	
370 375 380	
gag aga gtg ttt gat gga gag ctg caa gag gga cag gtg tta att gtg	1200
Glu Arg Val Phe Asp Gly Glu Leu Gln Glu Gly Gln Val Leu Ile Val	
385 390 395 400	
cca caa aac ttt gcg gtg gct gca aga tca cag agc gac aac ttc gag	1248
Pro Gln Asn Phe Ala Val Ala Ala Arg Ser Gln Ser Asp Asn Phe Glu	
405 410 415	

tat gtt tca ttc aag acc aat gat aga ccc tcg atc ggc aac ctt gca 1296  
 Tyr Val Ser Phe Lys Thr Asn Asp Arg Pro Ser Ile Gly Asn Leu Ala  
 420 425 430

ggt gca aac tca ttg ttg aac gca ttg ccg gag gaa gtg att cag caa 1344  
 Gly Ala Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu Val Ile Gln Gln  
 435 440 445

act ttt aac cta agg agg cag cag gcc agg cag gtc aag aac aac aac 1392  
 Thr Phe Asn Leu Arg Arg Gln Gln Ala Arg Gln Val Lys Asn Asn Asn  
 450 455 460

cct ttc agc ttc ctg gtt cca cct aag gag tct cag agg aga gtt gtg 1440  
 Pro Phe Ser Phe Leu Val Pro Pro Lys Glu Ser Gln Arg Arg Val Val  
 465 470 475 480

gct tag 1446  
 Ala

<210> 28  
 <211> 481  
 <212> PRT  
 <213> Glycine max

<400> 28

Met Ala Lys Leu Val Leu Ser Leu Cys Phe Leu Leu Phe Ser Gly Cys  
 1 5 10 15

Cys Phe Ala Phe Ser Phe Arg Glu Gln Pro Gln Gln Asn Glu Cys Gln  
 20 25 30

Ile Gln Arg Leu Asn Ala Leu Lys Pro Asp Asn Arg Ile Glu Ser Glu  
 35 40 45

Gly Gly Phe Ile Glu Thr Trp Asn Pro Asn Asn Lys Pro Phe Gln Cys  
 50 55 60

Ala Gly Val Ala Leu Ser Arg Cys Thr Leu Asn Arg Asn Ala Leu Arg  
 65 70 75 80

Arg Pro Ser Tyr Thr Asn Ala Pro Gln Glu Ile Tyr Ile Gln Gln Gly  
 85 90 95

Ser Gly Ile Phe Gly Met Ile Phe Pro Gly Cys Pro Ser Thr Phe Glu  
 100 105 110

Glu Pro Gln Gln Lys Gly Gln Ser Ser Arg Pro Gln Asp Arg His Gln  
 115 120 125

Lys Ile Tyr His Phe Arg Glu Gly Asp Leu Ile Ala Val Pro Thr Gly  
 130 135 140

Phe Ala Tyr Trp Met Tyr Asn Asn Glu Asp Thr Pro Val Val Ala Val  
 145 150 155 160

Ser Leu Ile Asp Thr Asn Ser Phe Gln Asn Gln Leu Asp Gln Met Pro  
 165 170 175

Arg Arg Phe Tyr Leu Ala Gly Asn Gln Glu Gln Glu Phe Leu Gln Tyr  
 180 185 190

Gln Pro Gln Lys Gln Gln Gly Gly Thr Gln Ser Gln Lys Gly Lys Arg  
 195 200 205

Gln Gln Glu Glu Glu Asn Glu Gly Gly Ser Ile Leu Ser Gly Phe Ala  
 210 215 220

Pro Glu Phe Leu Glu His Ala Phe Val Val Asp Arg Gln Ile Val Arg  
 225 230 235 240

## 39

Lys Leu Gln Gly Glu Asn Glu Glu Glu Glu Lys Gly Ala Ile Val Thr  
 245 250 255  
 Val Lys Gly Gly Leu Ser Val Ile Ser Pro Pro Thr Glu Glu Gln Gln  
 260 265 270  
 Gln Arg Pro Glu Glu Glu Glu Lys Pro Asp Cys Asp Glu Lys Asp Lys  
 275 280 285  
 His Cys Gln Ser Gln Ser Arg Asn Gly Ile Asp Glu Thr Ile Cys Thr  
 290 295 300  
 Met Arg Leu Arg His Asn Ile Gly Gln Thr Ser Ser Pro Asp Ile Phe  
 305 310 315 320  
 Asn Pro Gln Ala Gly Ser Ile Thr Thr Ala Thr Ser Leu Asp Phe Pro  
 325 330 335  
 Ala Leu Ser Trp Leu Lys Leu Ser Ala Gln Phe Gly Ser Leu Arg Lys  
 340 345 350  
 Asn Ala Met Phe Val Pro His Tyr Asn Leu Asn Ala Asn Ser Ile Ile  
 355 360 365  
 Tyr Ala Leu Asn Gly Arg Ala Leu Val Gln Val Val Asn Cys Asn Gly  
 370 375 380  
 Glu Arg Val Phe Asp Gly Glu Leu Gln Glu Gly Gln Val Leu Ile Val  
 385 390 395 400  
 Pro Gln Asn Phe Ala Val Ala Ala Arg Ser Gln Ser Asp Asn Phe Glu  
 405 410 415  
 Tyr Val Ser Phe Lys Thr Asn Asp Arg Pro Ser Ile Gly Asn Leu Ala  
 420 425 430  
 Gly Ala Asn Ser Leu Leu Asn Ala Leu Pro Glu Glu Val Ile Gln Gln  
 435 440 445  
 Thr Phe Asn Leu Arg Arg Gln Gln Ala Arg Gln Val Lys Asn Asn Asn  
 450 455 460  
 Pro Phe Ser Phe Leu Val Pro Pro Lys Glu Ser Gln Arg Arg Val Val  
 465 470 475 480  
 Ala

&lt;210&gt; 29

&lt;211&gt; 1482

&lt;212&gt; DNA

&lt;213&gt; Helianthus annuus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1479)

&lt;223&gt; 11S storage protein G3-D1

&lt;400&gt; 29

atg gca tcc aaa gca act ttg ctc tta gct ttt acc ctt ctc ttt gcc 48  
 Met Ala Ser Lys Ala Thr Leu Leu Leu Ala Phe Thr Leu Leu Phe Ala  
 1 5 10 15

act tgc att gcc cgc cac cag caa cgg caa cag caa cag aac cag tgc 96  
 Thr Cys Ile Ala Arg His Gln Gln Arg Gln Gln Gln Gln Asn Gln Cys  
 20 25 30

## 40

cag ctt caa aac atc gag gcg ctc gag ccc atc gaa gtt atc caa gct	144
Gln Leu Gln Asn Ile Glu Ala Leu Glu Pro Ile Glu Val Ile Gln Ala	
35 40 45	
gaa gcc ggt gtg acc gaa att tgg gac gcc tat gac caa cag ttc cag	192
Glu Ala Gly Val Thr Glu Ile Trp Asp Ala Tyr Asp Gln Gln Phe Gln	
50 55 60	
tgt gcg tgg tcg att tta ttc gac acc gga ttc aac ctg gtg gcc ttc	240
Cys Ala Trp Ser Ile Leu Phe Asp Thr Gly Phe Asn Leu Val Ala Phe	
65 70 75 80	
tct tgc ctt cct acg tca aca ccc cta ttt tgg cct tcg tcg aga gag	288
Ser Cys Leu Pro Thr Ser Thr Pro Leu Phe Trp Pro Ser Ser Arg Glu	
85 90 95	
ggg gtt ata ttg ccg gga tgc cgc aga acc tat gaa tat tcg cag gag	336
Gly Val Ile Leu Pro Gly Cys Arg Arg Thr Tyr Glu Tyr Ser Gln Glu	
100 105 110	
caa cag ttt tcc ggt gag ggt ggc cgc aga gga gga gga gag ggc aca	384
Gln Gln Phe Ser Gly Glu Gly Gly Arg Arg Gly Gly Glu Gly Thr	
115 120 125	
ttc agg acc gtc atc aga aag tta gag aac tta aag gag ggt gac gtg	432
Phe Arg Thr Val Ile Arg Lys Leu Glu Asn Leu Lys Glu Gly Asp Val	
130 135 140	
gtt gcc atc ccc acc gga aca gct cac tgg ctt cac aac gac ggc aac	480
Val Ala Ile Pro Thr Gly Thr Ala His Trp Leu His Asn Asp Gly Asn	
145 150 155 160	
aca gaa ctt gtg gtc gtc ttc ttg gat act cag aac cat gag aac cag	528
Thr Glu Leu Val Val Val Phe Leu Asp Thr Gln Asn His Glu Asn Gln	
165 170 175	
ctt gac gaa aac caa agg aga ttc ttc tta gcc gga aac cct caa gct	576
Leu Asp Glu Asn Gln Arg Arg Phe Phe Leu Ala Gly Asn Pro Gln Ala	
180 185 190	
caa gct caa agc cag cag caa caa caa aga caa cca cgc caa caa tct	624
Gln Ala Gln Ser Gln Gln Gln Gln Gln Arg Gln Pro Arg Gln Gln Ser	
195 200 205	
cct caa agg caa agg caa agg caa agg caa ggg caa ggt cag aac gcc	672
Pro Gln Arg Gln Arg Gln Arg Gln Arg Gln Gly Gln Gly Gln Asn Ala	
210 215 220	
ggc aac atc ttc aac ggt ttc acc ccc gag ctc att gca caa tca ttc	720
Gly Asn Ile Phe Asn Gly Phe Thr Pro Glu Leu Ile Ala Gln Ser Phe	
225 230 235 240	
aac gtc gac caa gag acc gcc cag aag cta caa gga caa aac gac cag	768
Asn Val Asp Gln Glu Thr Ala Gln Lys Leu Gln Gly Gln Asn Asp Gln	
245 250 255	
aga ggc cac att gtt aat gtc gga caa gac ctt caa ata gtc cgc cca	816
Arg Gly His Ile Val Asn Val Gly Gln Asp Leu Gln Ile Val Arg Pro	
260 265 270	
cca caa gac aga cgc tct cct cgc caa caa caa gag caa gcg acg tct	864
Pro Gln Asp Arg Arg Ser Pro Arg Gln Gln Gln Glu Gln Ala Thr Ser	
275 280 285	
cct cgc caa caa caa gag cag cag caa ggc aga cgt ggc gga tgg agc	912
Pro Arg Gln Gln Gln Glu Gln Gln Gln Gly Arg Arg Gly Gly Trp Ser	
290 295 300	

aac ggt gtg gaa gaa acc atc tgc agc atg aag ttc aaa gtg aac att 960  
 Asn Gly Val Glu Glu Thr Ile Cys Ser Met Lys Phe Lys Val Asn Ile  
 305 310 315 320  
 gac aac cct tcc cag gct gac ttt gta aac ccg caa gcc ggc agc att 1008  
 Asp Asn Pro Ser Gln Ala Asp Phe Val Asn Pro Gln Ala Gly Ser Ile  
 325 330 335  
 gca aac ctc aac agc ttc aaa ttc ccc att ctc gag cac ctc cgg ctc 1056  
 Ala Asn Leu Asn Ser Phe Lys Phe Pro Ile Leu Glu His Leu Arg Leu  
 340 345 350  
 agc gtg gaa aga ggc gaa ctc cgt ccg aat gcc atc caa tcc cca cac 1104  
 Ser Val Glu Arg Gly Glu Leu Arg Pro Asn Ala Ile Gln Ser Pro His  
 355 360 365  
 tgg aca atc aac gcc cac aat ctt ctc tac gta acc gag gga gcc ttg 1152  
 Trp Thr Ile Asn Ala His Asn Leu Leu Tyr Val Thr Glu Gly Ala Leu  
 370 375 380  
 agg gta caa atc gtc gac aac caa gga aac tca gtt ttc gac aac gag 1200  
 Arg Val Gln Ile Val Asp Asn Gln Gly Asn Ser Val Phe Asp Asn Glu  
 385 390 395 400  
 ctc cgt gag gga cag gtg gtg gtg atc ccg cag aac ttt gcg gtg atc 1248  
 Leu Arg Glu Gly Gln Val Val Val Ile Pro Gln Asn Phe Ala Val Ile  
 405 410 415  
 aag aga gcc aat gaa caa gga agc agg tgg gtg tct ttc aag act aat 1296  
 Lys Arg Ala Asn Glu Gln Gly Ser Arg Trp Val Ser Phe Lys Thr Asn  
 420 425 430  
 gat aat gcc atg ata gca aac ctt gca ggg cgt gtg tcc gca tca gca 1344  
 Asp Asn Ala Met Ile Ala Asn Leu Ala Gly Arg Val Ser Ala Ser Ala  
 435 440 445  
 gca tcg ccg ttg acg ttg tgg gcg aat ccg tat cag cta tct cga gag 1392  
 Ala Ser Pro Leu Thr Leu Trp Ala Asn Arg Tyr Gln Leu Ser Arg Glu  
 450 455 460  
 gaa gct cag cag ctc aag ttt agc cag agg gag acg gtt ttg ttt gca 1440  
 Glu Ala Gln Gln Leu Lys Phe Ser Gln Arg Glu Thr Val Leu Phe Ala  
 465 470 475 480  
 cca agt ttt tcc agg ggc caa ggg atc agg gct tca cgt taa 1482  
 Pro Ser Phe Ser Arg Gly Gln Gly Ile Arg Ala Ser Arg  
 485 490  
 <210> 30  
 <211> 493  
 <212> PRT  
 <213> Helianthus annuus  
 <400> 30  
 Met Ala Ser Lys Ala Thr Leu Leu Leu Ala Phe Thr Leu Leu Phe Ala  
 1 5 10 15  
 Thr Cys Ile Ala Arg His Gln Gln Arg Gln Gln Gln Gln Asn Gln Cys  
 20 25 30  
 Gln Leu Gln Asn Ile Glu Ala Leu Glu Pro Ile Glu Val Ile Gln Ala  
 35 40 45  
 Glu Ala Gly Val Thr Glu Ile Trp Asp Ala Tyr Asp Gln Gln Phe Gln  
 50 55 60  
 Cys Ala Trp Ser Ile Leu Phe Asp Thr Gly Phe Asn Leu Val Ala Phe  
 65 70 75 80



Ser	Cys	Leu	Pro	Thr	Ser	Thr	Pro	Leu	Phe	Trp	Pro	Ser	Ser	Arg	Glu	85	90	95
Gly	Val	Ile	Leu	Pro	Gly	Cys	Arg	Arg	Thr	Tyr	Glu	Tyr	Ser	Gln	Glu	100	105	110
Gln	Gln	Phe	Ser	Gly	Glu	Gly	Gly	Arg	Arg	Gly	Gly	Gly	Glu	Gly	Thr	115	120	125
Phe	Arg	Thr	Val	Ile	Arg	Lys	Leu	Glu	Asn	Leu	Lys	Glu	Gly	Asp	Val	130	135	140
Val	Ala	Ile	Pro	Thr	Gly	Thr	Ala	His	Trp	Leu	His	Asn	Asp	Gly	Asn	145	150	155
Thr	Glu	Leu	Val	Val	Val	Phe	Leu	Asp	Thr	Gln	Asn	His	Glu	Asn	Gln	165	170	175
Leu	Asp	Glu	Asn	Gln	Arg	Arg	Phe	Phe	Leu	Ala	Gly	Asn	Pro	Gln	Ala	180	185	190
Gln	Ala	Gln	Ser	Gln	Gln	Gln	Gln	Gln	Arg	Gln	Pro	Arg	Gln	Gln	Ser	195	200	205
Pro	Gln	Arg	Gln	Arg	Gln	Arg	Gln	Arg	Gln	Gly	Gln	Gly	Gln	Asn	Ala	210	215	220
Gly	Asn	Ile	Phe	Asn	Gly	Phe	Thr	Pro	Glu	Leu	Ile	Ala	Gln	Ser	Phe	225	230	235
Asn	Val	Asp	Gln	Glu	Thr	Ala	Gln	Lys	Leu	Gln	Gly	Gln	Asn	Asp	Gln	245	250	255
Arg	Gly	His	Ile	Val	Asn	Val	Gly	Gln	Asp	Leu	Gln	Ile	Val	Arg	Pro	260	265	270
Pro	Gln	Asp	Arg	Arg	Ser	Pro	Arg	Gln	Gln	Gln	Glu	Gln	Ala	Thr	Ser	275	280	285
Pro	Arg	Gln	Gln	Gln	Glu	Gln	Gln	Gln	Gly	Arg	Arg	Gly	Gly	Trp	Ser	290	295	300
Asn	Gly	Val	Glu	Glu	Thr	Ile	Cys	Ser	Met	Lys	Phe	Lys	Val	Asn	Ile	305	310	315
Asp	Asn	Pro	Ser	Gln	Ala	Asp	Phe	Val	Asn	Pro	Gln	Ala	Gly	Ser	Ile	325	330	335
Ala	Asn	Leu	Asn	Ser	Phe	Lys	Phe	Pro	Ile	Leu	Glu	His	Leu	Arg	Leu	340	345	350
Ser	Val	Glu	Arg	Gly	Glu	Leu	Arg	Pro	Asn	Ala	Ile	Gln	Ser	Pro	His	355	360	365
Trp	Thr	Ile	Asn	Ala	His	Asn	Leu	Leu	Tyr	Val	Thr	Glu	Gly	Ala	Leu	370	375	380
Arg	Val	Gln	Ile	Val	Asp	Asn	Gln	Gly	Asn	Ser	Val	Phe	Asp	Asn	Glu	385	390	395
Leu	Arg	Glu	Gly	Gln	Val	Val	Val	Ile	Pro	Gln	Asn	Phe	Ala	Val	Ile	405	410	415
Lys	Arg	Ala	Asn	Glu	Gln	Gly	Ser	Arg	Trp	Val	Ser	Phe	Lys	Thr	Asn	420	425	430
Asp	Asn	Ala	Met	Ile	Ala	Asn	Leu	Ala	Gly	Arg	Val	Ser	Ala	Ser	Ala	435	440	445
Ala	Ser	Pro	Leu	Thr	Leu	Trp	Ala	Asn	Arg	Tyr	Gln	Leu	Ser	Arg	Glu	450	455	460

Glu Ala Gln Gln Leu Lys Phe Ser Gln Arg Glu Thr Val Leu Phe Ala  
 465 470 475 480  
 Pro Ser Phe Ser Arg Gly Gln Gly Ile Arg Ala Ser Arg  
 485 490

<210> 31  
 <211> 537  
 <212> DNA  
 <213> Brassica napus

<220>  
 <221> CDS  
 <222> (1)..(534)  
 <223> NAPIN

<400> 31  
 atg gcg aac aag ctc ttc ctc gtc tcg gca act ctc gcc ttc ttc ttc 48  
 Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe  
 1 5 10 15  
 ctt ctc acc aat gcc tcc atc tac cgg acg gtc gtc gag ttc gac gaa 96  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu  
 20 25 30  
 gat gat gcc aca gac tca gcc ggc cca ttt agg att cca aaa tgt agg 144  
 Asp Asp Ala Thr Asp Ser Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg  
 35 40 45  
 aag gag ttt cag caa gca caa cac cta aga gct tgc cag cag tgg ctc 192  
 Lys Glu Phe Gln Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu  
 50 55 60  
 cac aag caa gca atg cag tct ggc ggt ggt cct agc tgg acc ctc gac 240  
 His Lys Gln Ala Met Gln Ser Gly Gly Gly Pro Ser Trp Thr Leu Asp  
 65 70 75 80  
 ggt gag ttt gac ttt gaa gac gac atg gag aac ccg cag ggt cca cag 288  
 Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln  
 85 90 95  
 cag aga ccg cct cta ctc cag cag tgc tgt aac gag ctc cac cag gaa 336  
 Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu  
 100 105 110  
 gag ccc ctt tgc gtt tgc cca acc ttg aaa gga gca tcc aaa gcg gtt 384  
 Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val  
 115 120 125  
 aaa caa caa att caa caa cag gga caa cag caa gga aag cag caa atg 432  
 Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met  
 130 135 140  
 gtg agc cgt atc tac cag acc gct acg cac tta cct aaa gtt tgc aac 480  
 Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn  
 145 150 155 160  
 atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct ggg ccc 528  
 Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro  
 165 170 175  
 tcc tac tag 537  
 Ser Tyr

<210> 32  
 <211> 178

&lt;212&gt; PRT

&lt;213&gt; Brassica napus

&lt;400&gt; 32

```

Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
  1              5              10              15
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
              20              25              30
Asp Asp Ala Thr Asp Ser Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
              35              40              45
Lys Glu Phe Gln Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu
  50              55              60
His Lys Gln Ala Met Gln Ser Gly Gly Gly Pro Ser Trp Thr Leu Asp
  65              70              75              80
Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln
              85              90              95
Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
              100             105             110
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
              115             120             125
Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met
              130             135             140
Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
              145             150             155             160
Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
              165             170             175
Ser Tyr

```

&lt;210&gt; 33

&lt;211&gt; 537

&lt;212&gt; DNA

&lt;213&gt; Brassica juncea

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(534)

&lt;223&gt; 2S storage protein

&lt;400&gt; 33

```

atg gcg aac aag ctc ttc ctc gtc tcg gca act ctc gcc ttc ttc ttc      48
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
  1              5              10              15
ctt ctc acc aat gcc tcc atc tac cgg acg gtc gtc gag ttc gac gaa      96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
              20              25              30
gat gat gcc aca gac tca gcc ggc cca ttt agg att cca aaa tgt agg     144
Asp Asp Ala Thr Asp Ser Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
              35              40              45
aag gag ttt cag caa gca caa cac cta aga gtc tgc cag cag tgg ctc     192
Lys Glu Phe Gln Gln Ala Gln His Leu Arg Val Cys Gln Gln Trp Leu
  50              55              60

```

## 45

```

cac aag caa gca atg cag tct ggc ggt ggt ctc agc tgg acc ctc gac      240
His Lys Gln Ala Met Gln Ser Gly Gly Gly Leu Ser Trp Thr Leu Asp
  65                               70                               75                               80

ggt gag ttt gac ttt gaa gac gac atg gag aac tcg cag ggt cca cag      288
Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Ser Gln Gly Pro Gln
                               85                               90                               95

cag aga ccg cct cta ctc cag cag tgc tgt aac gag ctc cac cag gaa      336
Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
                               100                              105                              110

gag ccc ctt tgc gtt tgc cca acc ttg aaa gga gca tcc aaa gcg gtt      384
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
                               115                              120                              125

aaa caa caa att caa caa cag gga caa cag caa gga aag cag caa atg      432
Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met
                               130                              135                              140

gtg agc cgt atc tac cag acc gct acg cac tta cct aaa gtt tgc aac      480
Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
                               145                              150                              155                              160

atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct ggg ccc      528
Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
                               165                              170                              175

tcc tac tag                                                                537
Ser Tyr

<210> 34
<211> 178
<212> PRT
<213> Brassica juncea

<400> 34
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
  1                               5                               10                               15

Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
                               20                               25                               30

Asp Asp Ala Thr Asp Ser Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
                               35                               40                               45

Lys Glu Phe Gln Gln Ala Gln His Leu Arg Val Cys Gln Gln Trp Leu
                               50                               55                               60

His Lys Gln Ala Met Gln Ser Gly Gly Gly Leu Ser Trp Thr Leu Asp
  65                               70                               75                               80

Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Ser Gln Gly Pro Gln
                               85                               90                               95

Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
                               100                              105                              110

Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
                               115                              120                              125

Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met
  130                              135                              140

Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
  145                              150                              155                              160

Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
                               165                              170                              175

```

Ser Tyr

&lt;210&gt; 35

&lt;211&gt; 537

&lt;212&gt; DNA

&lt;213&gt; Brassica oleracea

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(534)

&lt;223&gt; 2S storage protein

&lt;400&gt; 35

```

atg gcg aac aag ctc ttc ctc gtc tcg gca act ctc gcc ttc ttc ttc      48
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
   1             5             10             15

ctt ctc acc aat gcc tcc atc tac cgg acg gtg gtc gag ttc gac gaa      96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
             20             25             30

gat gat gcc aca aac cca gcc ggc cca ttt agg atc cca aaa tgt agg     144
Asp Asp Ala Thr Asn Pro Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
             35             40             45

aag gag ttt cag caa gca caa cac cta aga gct tgc cag cag tgg ctc     192
Lys Glu Phe Gln Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu
             50             55             60

cac aag caa gca atg cag tct ggc ggt ggt cct agc tgg acc ctc gac     240
His Lys Gln Ala Met Gln Ser Gly Gly Gly Pro Ser Trp Thr Leu Asp
             65             70             75             80

agt gag ttt gac ttt gaa gac gac atg gag aac ccg cag ggt cca cag     288
Ser Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln
             85             90             95

cag aga ccg cct cta ctc ctg caa tgc tgt aac gag ctg gac cag gaa     336
Gln Arg Pro Pro Leu Leu Leu Gln Cys Cys Asn Glu Leu Asp Gln Glu
             100            105            110

gag ccc ctt tgc gtt tgc cca acc ttg aaa gga gca tcc aaa gcg gtt     384
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
             115            120            125

aaa caa caa att caa caa cag gga caa cag caa gga aag cag caa atg     432
Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met
             130            135            140

gtg agc cgt atc tac cag acc gct acg cac tta cct aaa gtt tgc aac     480
Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
             145            150            155            160

atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct ggg ccc     528
Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
             165            170            175

tcc tac tag
Ser Tyr

```

&lt;210&gt; 36

&lt;211&gt; 178

&lt;212&gt; PRT

&lt;213&gt; Brassica oleracea

&lt;400&gt; 36

```

Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
 1           5           10           15
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu
           20           25           30
Asp Asp Ala Thr Asn Pro Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
           35           40           45
Lys Glu Phe Gln Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu
           50           55           60
His Lys Gln Ala Met Gln Ser Gly Gly Gly Pro Ser Trp Thr Leu Asp
           65           70           75           80
Ser Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln
           85           90           95
Gln Arg Pro Pro Leu Leu Leu Gln Cys Cys Asn Glu Leu Asp Gln Glu
           100          105          110
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
           115          120          125
Lys Gln Gln Ile Gln Gln Gln Gly Gln Gln Gln Gly Lys Gln Gln Met
           130          135          140
Val Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
           145          150          155          160
Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro
           165          170          175
Ser Tyr

```

&lt;210&gt; 37

&lt;211&gt; 543

&lt;212&gt; DNA

&lt;213&gt; Brassica napus cv. Topas

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(540)

&lt;223&gt; Napin

&lt;400&gt; 37

```

atg gcg aac aag ctc ttc ctc gtc tcg gca act ctt gcc ttc ttc ttc      48
Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
 1           5           10           15
ctt ctc acc aac gcc tcc atc tac cgc acc atc gtg gaa gtc gac gaa      96
Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Ile Val Glu Val Asp Glu
           20           25           30
gat gat gcc aca aac cca gcc ggc cca ttt agg att cca aaa tgt agg      144
Asp Asp Ala Thr Asn Pro Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
           35           40           45
aag gag ttt cag caa gca caa cac ctg aaa gct tgc caa caa tgg ctc      192
Lys Glu Phe Gln Gln Ala Gln His Leu Lys Ala Cys Gln Gln Trp Leu
           50           55           60
cac aag cag gca atg cag tcc ggt agt ggc cca agc tgg acc ctc gac      240
His Lys Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp
           65           70           75           80

```

48

```

ggt gag ttt gat ttt gaa gat gac atg gag aac ccc cag ggc cca caa 288
Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln
      85                      90                      95

cag agg ccg cca cta ctc cag cag tgc tgc aac gag ctc cac cag gaa 336
Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
      100                      105                      110

gag cca ctt tgc gtt tgc cca acc ttg aaa gga gca tcc aaa gcc gtt 384
Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
      115                      120                      125

aaa caa cag gtt cga caa cag caa gga cag cag gga cag cag ctg cag 432
Lys Gln Gln Val Arg Gln Gln Gln Gly Gln Gln Gly Gln Gln Leu Gln
      130                      135                      140

caa gta att agc cgt atc tac cag act gct acg cac tta cct aaa gtt 480
Gln Val Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val
      145                      150                      155                      160

tgc aac atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct 528
Cys Asn Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro
      165                      170                      175

gga ccc tcc tac tag 543
Gly Pro Ser Tyr
      180

```

&lt;210&gt; 38

&lt;211&gt; 180

&lt;212&gt; PRT

&lt;213&gt; Brassica napus cv. Topas

&lt;400&gt; 38

```

Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe
  1                      5                      10                      15

Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Ile Val Glu Val Asp Glu
      20                      25                      30

Asp Asp Ala Thr Asn Pro Ala Gly Pro Phe Arg Ile Pro Lys Cys Arg
      35                      40                      45

Lys Glu Phe Gln Gln Ala Gln His Leu Lys Ala Cys Gln Gln Trp Leu
      50                      55                      60

His Lys Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp
      65                      70                      75                      80

Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln
      85                      90                      95

Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu
      100                      105                      110

Glu Pro Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val
      115                      120                      125

Lys Gln Gln Val Arg Gln Gln Gln Gly Gln Gln Gly Gln Gln Leu Gln
      130                      135                      140

Gln Val Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val
      145                      150                      155                      160

Cys Asn Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro
      165                      170                      175

Gly Pro Ser Tyr
      180

```

<210> 39  
 <211> 435  
 <212> DNA  
 <213> Sinapis alba

<220>  
 <221> CDS  
 <222> (1)..(432)  
 <223> coding for partial sin1 storage protein

<400> 39

cca gcc ggc cca ttt ggg att cca aaa tgt agg aag gag ttt caa caa	48
Pro Ala Gly Pro Phe Gly Ile Pro Lys Cys Arg Lys Glu Phe Gln Gln	
1 5 10 15	
gca caa cac cta aga gct tgc cag caa tgg ctc cac aag cag gca atg	96
Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu His Lys Gln Ala Met	
20 25 30	
cag tct ggt agt ggt cca agc tgg acc ctc gac gat gag ttt gat ttt	144
Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp Asp Glu Phe Asp Phe	
35 40 45	
gaa gat gac atg gag aac cca cag gga cca cag cag agg cca cca cta	192
Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln Gln Arg Pro Pro Leu	
50 55 60	
ctc cag cag tgc tgc aac gag ctc cac cag gaa gag cca ctt tgc gtt	240
Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu Glu Pro Leu Cys Val	
65 70 75 80	
tgc cca acc ttg aaa gga gca tcc aaa gcc gtt aaa cag cag gtt aga	288
Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val Lys Gln Gln Val Arg	
85 90 95	
caa cag ctg gag cag cag gga cag cag gga ccg cac ctg cag cat gta	336
Gln Gln Leu Glu Gln Gln Gly Gln Gln Gly Pro His Leu Gln His Val	
100 105 110	
att agc cgt atc tac cag acc gct acg cac tta cct aga gtt tgc aac	384
Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Arg Val Cys Asn	
115 120 125	
att agg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct gga ccc	432
Ile Arg Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro	
130 135 140	
tcc	435

<210> 40  
 <211> 144  
 <212> PRT  
 <213> Sinapis alba

<400> 40

Pro Ala Gly Pro Phe Gly Ile Pro Lys Cys Arg Lys Glu Phe Gln Gln
1 5 10 15
Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu His Lys Gln Ala Met
20 25 30
Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp Asp Glu Phe Asp Phe
35 40 45
Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln Gln Arg Pro Pro Leu
50 55 60



## 50

Leu	Gln	Gln	Cys	Cys	Asn	Glu	Leu	His	Gln	Glu	Glu	Pro	Leu	Cys	Val	
65					70					75					80	
Cys	Pro	Thr	Leu	Lys	Gly	Ala	Ser	Lys	Ala	Val	Lys	Gln	Gln	Val	Arg	
				85					90					95		
Gln	Gln	Leu	Glu	Gln	Gln	Gly	Gln	Gln	Gly	Pro	His	Leu	Gln	His	Val	
			100					105					110			
Ile	Ser	Arg	Ile	Tyr	Gln	Thr	Ala	Thr	His	Leu	Pro	Arg	Val	Cys	Asn	
		115					120					125				
Ile	Arg	Gln	Val	Ser	Val	Cys	Pro	Phe	Gln	Lys	Thr	Met	Pro	Gly	Pro	
		130				135						140				

&lt;210&gt; 41

&lt;211&gt; 468

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(465)

&lt;223&gt; 2S albumine 1

&lt;400&gt; 41

atg	acc	aag	ctt	aca	att	ctc	ctc	atc	gct	ctt	ctc	ttc	atc	gcc	cac	48
Met	Thr	Lys	Leu	Thr	Ile	Leu	Leu	Ile	Ala	Leu	Leu	Phe	Ile	Ala	His	
1				5					10					15		
acc	tgc	tgc	gcc	tcc	aaa	tgg	caa	cag	cac	cag	caa	gag	agc	tgc	cgc	96
Thr	Cys	Cys	Ala	Ser	Lys	Trp	Gln	Gln	His	Gln	Gln	Glu	Ser	Cys	Arg	
			20					25					30			
gag	cag	ctc	aag	ggg	atc	aac	ctc	aac	ccc	tgt	gag	cac	atc	atg	gag	144
Glu	Gln	Leu	Lys	Gly	Ile	Asn	Leu	Asn	Pro	Cys	Glu	His	Ile	Met	Glu	
			35				40					45				
aag	atc	caa	gct	ggc	cgc	cgc	ggc	gag	gac	ggc	agc	gac	gaa	gat	cac	192
Lys	Ile	Gln	Ala	Gly	Arg	Arg	Gly	Glu	Asp	Gly	Ser	Asp	Glu	Asp	His	
	50					55					60					
att	ctc	atc	agg	acc	atg	ccg	gga	aga	atc	aac	tac	atc	agg	aag	aag	240
Ile	Leu	Ile	Arg	Thr	Met	Pro	Gly	Arg	Ile	Asn	Tyr	Ile	Arg	Lys	Lys	
	65				70					75					80	
gaa	gga	aaa	gaa	gaa	gaa	gaa	gaa	gga	cac	atg	cag	aag	tgc	tgc	agc	288
Glu	Gly	Lys	Glu	Glu	Glu	Glu	Glu	Gly	His	Met	Gln	Lys	Cys	Cys	Ser	
				85					90					95		
gaa	atg	agc	gag	ctg	aaa	agc	ccc	ata	tgc	cag	tgc	aaa	gcg	cta	cag	336
Glu	Met	Ser	Glu	Leu	Lys	Ser	Pro	Ile	Cys	Gln	Cys	Lys	Ala	Leu	Gln	
			100					105					110			
aag	ata	atg	gat	aac	cag	agc	gag	caa	ctg	gag	ggg	aag	gag	aag	aag	384
Lys	Ile	Met	Asp	Asn	Gln	Ser	Glu	Gln	Leu	Glu	Gly	Lys	Glu	Lys	Lys	
		115					120					125				
cag	atg	gag	aga	gag	ctc	atg	aac	ttg	gct	att	agg	tgc	agg	ttg	gga	432
Gln	Met	Glu	Arg	Glu	Leu	Met	Asn	Leu	Ala	Ile	Arg	Cys	Arg	Leu	Gly	
		130				135					140					
ccc	atg	ata	ggg	tgc	gac	ttg	tcc	tcc	gat	gac	tga					468
Pro	Met	Ile	Gly	Cys	Asp	Leu	Ser	Ser	Asp	Asp						
					150					155						
145																

<210> 42  
 <211> 155  
 <212> PRT  
 <213> Glycine max

<400> 42

Met	Thr	Lys	Leu	Thr	Ile	Leu	Leu	Ile	Ala	Leu	Leu	Phe	Ile	Ala	His	
1				5					10					15		
Thr	Cys	Cys	Ala	Ser	Lys	Trp	Gln	Gln	His	Gln	Gln	Glu	Ser	Cys	Arg	
			20				25						30			
Glu	Gln	Leu	Lys	Gly	Ile	Asn	Leu	Asn	Pro	Cys	Glu	His	Ile	Met	Glu	
		35				40						45				
Lys	Ile	Gln	Ala	Gly	Arg	Arg	Gly	Glu	Asp	Gly	Ser	Asp	Glu	Asp	His	
	50					55					60					
Ile	Leu	Ile	Arg	Thr	Met	Pro	Gly	Arg	Ile	Asn	Tyr	Ile	Arg	Lys	Lys	
65					70					75					80	
Glu	Gly	Lys	Glu	Glu	Glu	Glu	Gly	His	Met	Gln	Lys	Cys	Cys	Ser		
			85				90						95			
Glu	Met	Ser	Glu	Leu	Lys	Ser	Pro	Ile	Cys	Gln	Cys	Lys	Ala	Leu	Gln	
			100				105						110			
Lys	Ile	Met	Asp	Asn	Gln	Ser	Glu	Gln	Leu	Glu	Gly	Lys	Glu	Lys	Lys	
		115					120					125				
Gln	Met	Glu	Arg	Glu	Leu	Met	Asn	Leu	Ala	Ile	Arg	Cys	Arg	Leu	Gly	
	130					135					140					
Pro	Met	Ile	Gly	Cys	Asp	Leu	Ser	Ser	Asp	Asp						
145					150					155						

<210> 43  
 <211> 477  
 <212> DNA  
 <213> Glycine max

<220>

<221> CDS

<222> (1)..(474)

<223> 2S albumine

<400> 43

atg	acc	aag	ttc	aca	atc	ctc	ctc	atc	tct	ctt	ctc	ttc	tgc	atc	gcc	48
Met	Thr	Lys	Phe	Thr	Ile	Leu	Leu	Ile	Ser	Leu	Leu	Phe	Cys	Ile	Ala	
1				5					10					15		
cac	act	tgc	agc	gcc	tcc	aaa	tgg	cag	cac	cag	caa	gat	agc	tgc	cgc	96
His	Thr	Cys	Ser	Ala	Ser	Lys	Trp	Gln	His	Gln	Gln	Asp	Ser	Cys	Arg	
			20					25					30			
aag	cag	ctc	cag	ggg	gtg	aac	ctc	acg	ccc	tgc	gag	aag	cac	atc	atg	144
Lys	Gln	Leu	Gln	Gly	Val	Asn	Leu	Thr	Pro	Cys	Glu	Lys	His	Ile	Met	
		35				40						45				
gag	aag	atc	caa	ggc	cgc	ggc	gat	gac	gat	gat	gat	gat	gac	gac	gac	192
Glu	Lys	Ile	Gln	Gly	Arg	Gly	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Asp	Asp	
		50				55						60				
aat	cac	att	ctc	agg	acc	atg	cgg	gga	aga	atc	aac	tac	ata	agg	agg	240
Asn	His	Ile	Leu	Arg	Thr	Met	Arg	Gly	Arg	Ile	Asn	Tyr	Ile	Arg	Arg	
65					70					75					80	

## 52

aac gaa gga aaa gac gaa gac gaa gaa gaa gaa gga cac atg cag aag 288  
 Asn Glu Gly Lys Asp Glu Asp Glu Glu Glu Glu Gly His Met Gln Lys  
                                     85                                    90                                    95  
 tgc tgc aca gaa atg agc gag ctg aga agc ccc aaa tgc cag tgc aaa 336  
 Cys Cys Thr Glu Met Ser Glu Leu Arg Ser Pro Lys Cys Gln Cys Lys  
                                     100                                    105                                    110  
 gcg ctg cag aag ata atg gag aac cag agc gag gaa ctg gag gag aag 384  
 Ala Leu Gln Lys Ile Met Glu Asn Gln Ser Glu Glu Leu Glu Glu Lys  
                                     115                                    120                                    125  
 cag aag aag aaa atg gag aag gag ctc att aac ttg gct act atg tgc 432  
 Gln Lys Lys Lys Met Glu Lys Glu Leu Ile Asn Leu Ala Thr Met Cys  
                                     130                                    135                                    140  
 agg ttt gga ccc atg atc cag tgc gac ttg tcc tcc gat gac taa 477  
 Arg Phe Gly Pro Met Ile Gln Cys Asp Leu Ser Ser Asp Asp  
                                     145                                    150                                    155

&lt;210&gt; 44

&lt;211&gt; 158

&lt;212&gt; PRT

&lt;213&gt; Glycine max

&lt;400&gt; 44

Met Thr Lys Phe Thr Ile Leu Leu Ile Ser Leu Leu Phe Cys Ile Ala  
   1                                    5                                    10                                    15  
 His Thr Cys Ser Ala Ser Lys Trp Gln His Gln Gln Asp Ser Cys Arg  
                                     20                                    25                                    30  
 Lys Gln Leu Gln Gly Val Asn Leu Thr Pro Cys Glu Lys His Ile Met  
                                     35                                    40                                    45  
 Glu Lys Ile Gln Gly Arg Gly Asp Asp Asp Asp Asp Asp Asp Asp  
                                     50                                    55                                    60  
 Asn His Ile Leu Arg Thr Met Arg Gly Arg Ile Asn Tyr Ile Arg Arg  
                                     65                                    70                                    75                                    80  
 Asn Glu Gly Lys Asp Glu Asp Glu Glu Glu Glu Gly His Met Gln Lys  
                                     85                                    90                                    95  
 Cys Cys Thr Glu Met Ser Glu Leu Arg Ser Pro Lys Cys Gln Cys Lys  
                                     100                                    105                                    110  
 Ala Leu Gln Lys Ile Met Glu Asn Gln Ser Glu Glu Leu Glu Glu Lys  
                                     115                                    120                                    125  
 Gln Lys Lys Lys Met Glu Lys Glu Leu Ile Asn Leu Ala Thr Met Cys  
                                     130                                    135                                    140  
 Arg Phe Gly Pro Met Ile Gln Cys Asp Leu Ser Ser Asp Asp  
                                     145                                    150                                    155

&lt;210&gt; 45

&lt;211&gt; 537

&lt;212&gt; DNA

&lt;213&gt; Brassica nigra

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(534)

&lt;223&gt; 2S storage protein

<400> 45  
 atg gcg aac aag ctc ttc ctc gtc tcg gca act ctc gcc ttc ttc ttc 48  
 Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe  
 1 5 10 15  
 ctg ctc acc aat gcc tcc atc tac cgg acg gtc gtc gag ttc gac gaa 96  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu  
 20 25 30  
 gat gat gac aca aac caa gcc gga cca ttt agg att cca aga tgt cga 144  
 Asp Asp Asp Thr Asn Gln Ala Gly Pro Phe Arg Ile Pro Arg Cys Arg  
 35 40 45  
 aag gag ttt cgg caa gca caa cat cta aga gct tgc cag caa tgg ctc 192  
 Lys Glu Phe Arg Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu  
 50 55 60  
 cac agg cag gca atg cag tcc ggt agt ggt cca agc tgg acc ctg gac 240  
 His Arg Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp  
 65 70 75 80  
 ggt gag ttt gac ttt gaa gac gac atg gag aac caa cag ggc cca cag 288  
 Gly Glu Phe Asp Phe Glu Asp Asp Met Glu Asn Gln Gln Gly Pro Gln  
 85 90 95  
 cag agg cca cct cta ctc cag caa tgc tgc aac gag ctc cac cag gaa 336  
 Gln Arg Pro Pro Leu Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu  
 100 105 110  
 gag gca ctt tgt gtt tgc cca acc ttg aaa gga gca tcc aaa gcg gtt 384  
 Glu Ala Leu Cys Val Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val  
 115 120 125  
 aga caa cag gtt cga caa cag gga cac cag cag cag atg cag cat gta 432  
 Arg Gln Gln Val Arg Gln Gln Gly His Gln Gln Met Gln His Val  
 130 135 140  
 att agc cgt atc tac cag acc gct acg cac tta cct aga gtt tgc aac 480  
 Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Arg Val Cys Asn  
 145 150 155 160  
 atc ccg caa gtt agc gtt tgt ccc ttc cag aag acc atg cct ggg ccc 528  
 Ile Pro Gln Val Ser Val Cys Pro Phe Gln Lys Thr Met Pro Gly Pro  
 165 170 175  
 tcc tac tag 537  
 Ser Tyr  
 <210> 46  
 <211> 178  
 <212> PRT  
 <213> Brassica nigra  
 <400> 46  
 Met Ala Asn Lys Leu Phe Leu Val Ser Ala Thr Leu Ala Phe Phe Phe  
 1 5 10 15  
 Leu Leu Thr Asn Ala Ser Ile Tyr Arg Thr Val Val Glu Phe Asp Glu  
 20 25 30  
 Asp Asp Asp Thr Asn Gln Ala Gly Pro Phe Arg Ile Pro Arg Cys Arg  
 35 40 45  
 Lys Glu Phe Arg Gln Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu  
 50 55 60  
 His Arg Gln Ala Met Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp  
 65 70 75 80

[illegible]

```
<210> 47
<211> 435
<212> DNA
<213> Sinapis alba

<220>
<221> CDS
<222> (1)..(432)
<223> sin5 storage protein
```

BNSDOCID: &lt;WO 0307B629A1 | &gt;

tcc

435

&lt;210&gt; 48

&lt;211&gt; 144

&lt;212&gt; PRT

&lt;213&gt; Sinapis alba

&lt;400&gt; 48

```

Pro Ala Gly Pro Phe Gly Ile Pro Lys Cys Arg Lys Glu Phe Gln Gln
 1           5           10           15
Ala Gln His Leu Arg Ala Cys Gln Gln Trp Leu His Lys Gln Ala Met
          20           25           30
Gln Ser Gly Ser Gly Pro Ser Trp Thr Leu Asp Asp Glu Phe Asp Phe
          35           40           45
Glu Asp Asp Met Glu Asn Pro Gln Gly Pro Gln Gln Lys Pro Pro Leu
          50           55           60
Leu Gln Gln Cys Cys Asn Glu Leu His Gln Glu Glu Pro Leu Cys Val
          65           70           75           80
Cys Pro Thr Leu Lys Gly Ala Ser Lys Ala Val Lys Gln Gln Val Arg
          85           90           95
Gln Gln Leu Gly Gln Gln Gly Gln Gln Gly Pro Gln Val Gln His Val
          100          105          110
Ile Ser Arg Ile Tyr Gln Thr Ala Thr His Leu Pro Lys Val Cys Asn
          115          120          125
Ile Pro Gln Val Ser Val Cys Pro Phe Lys Lys Thr Met Pro Gly Pro
          130          135          140

```

&lt;210&gt; 49

&lt;211&gt; 888

&lt;212&gt; DNA

&lt;213&gt; Helianthus annuus

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(885)

&lt;223&gt; HaG5 2S albumine

&lt;400&gt; 49

```

atg gca aag caa ata gtt ctc gca ctc gct ttc gcc gcc ctt gta gcc
Met Ala Lys Gln Ile Val Leu Ala Leu Ala Phe Ala Ala Leu Val Ala
 1           5           10           15
ttt gct acc gcc cac aca acc ata atc acc acc acc atc gaa gac gag
Phe Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Glu Asp Glu
          20           25           30
aac ccg atc tcc gga caa agg caa gtg agc caa cgg ata cag gga caa
Asn Pro Ile Ser Gly Gln Arg Gln Val Ser Gln Arg Ile Gln Gly Gln
          35           40           45
agg ctg aac cag tgt cgc atg ttc ctc cag cag ggt cag aac att cct
Arg Leu Asn Gln Cys Arg Met Phe Leu Gln Gln Gly Gln Asn Ile Pro
          50           55           60
cgc gaa ttc gat aac cct cag atg ggg cgg cag cag gag cag cag ctc
Arg Glu Phe Asp Asn Pro Gln Met Gly Arg Gln Gln Glu Gln Gln Leu
          65           70           75           80

```

## 56

cag cag tgt tgt caa gag ctc caa aac atc gaa ggg cag tgc caa tgt 288  
 Gln Gln Cys Cys Gln Glu Leu Gln Asn Ile Glu Gly Gln Cys Gln Cys  
                                     85                                    90                                    95

gag gcg gtg aag cag gtg ttc cga gaa gcc cag cag caa gta caa cag 336  
 Glu Ala Val Lys Gln Val Phe Arg Glu Ala Gln Gln Gln Val Gln Gln  
                                     100                                    105                                    110

caa cag gga cgg cag ctt gta ccc ttc cgc ggt tgc cag cag acc caa 384  
 Gln Gln Gly Arg Gln Leu Val Pro Phe Arg Gly Ser Gln Gln Thr Gln  
                                     115                                    120                                    125

cag ttg aag cag aag gct cag att ctc cct aac gta tgc aac ctt caa 432  
 Gln Leu Lys Gln Lys Ala Gln Ile Leu Pro Asn Val Cys Asn Leu Gln  
                                     130                                    135                                    140

tca aga cga tgt gaa atc gga acc atc acc acc acc gtc acc gag agc 480  
 Ser Arg Arg Cys Glu Ile Gly Thr Ile Thr Thr Thr Val Thr Glu Ser  
 145                                    150                                    155                                    160

aat atc gat atc ccc ttc cgt gac agg ccc ttt ggc act gga tca caa 528  
 Asn Ile Asp Ile Pro Phe Arg Asp Arg Pro Phe Gly Thr Gly Ser Gln  
                                     165                                    170                                    175

cag tgc aga gaa act gaa atc caa cga ccc gtt ggt gaa tgc caa agg 576  
 Gln Cys Arg Glu Thr Glu Ile Gln Arg Pro Val Gly Glu Cys Gln Arg  
                                     180                                    185                                    190

ttc gtg gag cag caa atg cag cag tct ccg agg tcc act aga cca tac 624  
 Phe Val Glu Gln Gln Met Gln Gln Ser Pro Arg Ser Thr Arg Pro Tyr  
                                     195                                    200                                    205

caa cag cgg cca gga caa cag cag cag cag cag aga ggg ctc caa caa 672  
 Gln Gln Arg Pro Gly Gln Gln Gln Gln Gln Gln Arg Gly Leu Gln Gln  
                                     210                                    215                                    220

caa tgc tgc aac gag cta caa aac gtg aag agg gag tgt cat tgc gag 720  
 Gln Cys Cys Asn Glu Leu Gln Asn Val Lys Arg Glu Cys His Cys Glu  
 225                                    230                                    235                                    240

gca att caa gaa gtg gct agg aga gtg atg agg cag cca cag cag cag 768  
 Ala Ile Gln Glu Val Ala Arg Arg Val Met Arg Gln Pro Gln Gln Gln  
                                     245                                    250                                    255

cag cag caa cgt cgt ggg cag ttc ggt ggg cag gag atg gaa acc gcg 816  
 Gln Gln Gln Arg Arg Gly Gln Phe Gly Gly Gln Glu Met Glu Thr Ala  
                                     260                                    265                                    270

agg agg gtg att cag aat ctg ccc aac cag tgc gac ttg gaa gtc cag 864  
 Arg Arg Val Ile Gln Asn Leu Pro Asn Gln Cys Asp Leu Glu Val Gln  
                                     275                                    280                                    285

caa tgc aca acc tgt acg gga tga 888  
 Gln Cys Thr Thr Cys Thr Gly  
                                     290                                    295

<210> 50  
 <211> 295  
 <212> PRT  
 <213> Helianthus annuus  
 <400> 50  
 Met Ala Lys Gln Ile Val Leu Ala Leu Ala Phe Ala Ala Leu Val Ala  
   1                                    5                                    10                                    15  
 Phe Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Glu Asp Glu  
                                     20                                    25                                    30

Asn	Pro	Ile	Ser	Gly	Gln	Arg	Gln	Val	Ser	Gln	Arg	Ile	Gln	Gly	Gln
		35						40				45			
Arg	Leu	Asn	Gln	Cys	Arg	Met	Phe	Leu	Gln	Gln	Gly	Gln	Asn	Ile	Pro
	50					55					60				
Arg	Glu	Phe	Asp	Asn	Pro	Gln	Met	Gly	Arg	Gln	Gln	Glu	Gln	Gln	Leu
65					70					75					80
Gln	Gln	Cys	Cys	Gln	Glu	Leu	Gln	Asn	Ile	Glu	Gly	Gln	Cys	Gln	Cys
				85					90					95	
Glu	Ala	Val	Lys	Gln	Val	Phe	Arg	Glu	Ala	Gln	Gln	Gln	Val	Gln	Gln
			100					105					110		
Gln	Gln	Gly	Arg	Gln	Leu	Val	Pro	Phe	Arg	Gly	Ser	Gln	Gln	Thr	Gln
		115					120					125			
Gln	Leu	Lys	Gln	Lys	Ala	Gln	Ile	Leu	Pro	Asn	Val	Cys	Asn	Leu	Gln
130						135					140				
Ser	Arg	Arg	Cys	Glu	Ile	Gly	Thr	Ile	Thr	Thr	Thr	Val	Thr	Glu	Ser
145					150					155					160
Asn	Ile	Asp	Ile	Pro	Phe	Arg	Asp	Arg	Pro	Phe	Gly	Thr	Gly	Ser	Gln
				165					170					175	
Gln	Cys	Arg	Glu	Thr	Glu	Ile	Gln	Arg	Pro	Val	Gly	Glu	Cys	Gln	Arg
			180					185					190		
Phe	Val	Glu	Gln	Gln	Met	Gln	Gln	Ser	Pro	Arg	Ser	Thr	Arg	Pro	Tyr
	195						200					205			
Gln	Gln	Arg	Pro	Gly	Gln	Gln	Gln	Gln	Gln	Gln	Arg	Gly	Leu	Gln	Gln
210						215					220				
Gln	Cys	Cys	Asn	Glu	Leu	Gln	Asn	Val	Lys	Arg	Glu	Cys	His	Cys	Glu
225					230					235					240
Ala	Ile	Gln	Glu	Val	Ala	Arg	Arg	Val	Met	Arg	Gln	Pro	Gln	Gln	Gln
				245					250					255	
Gln	Gln	Gln	Arg	Arg	Gly	Gln	Phe	Gly	Gly	Gln	Glu	Met	Glu	Thr	Ala
			260					265					270		
Arg	Arg	Val	Ile	Gln	Asn	Leu	Pro	Asn	Gln	Cys	Asp	Leu	Glu	Val	Gln
		275					280					285			
Gln	Cys	Thr	Thr	Cys	Thr	Gly									
290						295									

gct aca gcc cac aca acc atc atc acc acc acc atc gac gac gag aac 97  
Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Asp Asp Glu Asn  
20 25 30



ccg atc tcc gaa caa agg caa tgt tgg caa cag gta cag gga caa agg	145
Pro Ile Ser Glu Gln Arg Gln Cys Trp Gln Gln Val Gln Gly Gln Arg	
35 40 45	
ttg aac cag tgt cgc atg ttc ctc cag caa ggt cag agg ggg cag caa	193
Leu Asn Gln Cys Arg Met Phe Leu Gln Gln Gly Gln Arg Gly Gln Gln	
50 55 60	
cac caa cag caa cag cat cag cag cag gag cag cag ctc ctc cag cag	241
His Gln Gln Gln Gln His Gln Gln Gln Glu Gln Gln Leu Leu Gln Gln	
65 70 75 80	
tgt tgt caa gag ctt caa aac atc gaa gga cag tgc caa tgt gag gcg	289
Cys Cys Gln Glu Leu Gln Asn Ile Glu Gly Gln Cys Gln Cys Glu Ala	
85 90 95	
gtg aag cag gtg gtc cga gat gct cag cga cac gag caa cag cga ccg	337
Val Lys Gln Val Val Arg Asp Ala Gln Arg His Glu Gln Gln Arg Pro	
100 105 110	
cga gtg ccc ttc cag ggt tct cag cag tct caa cag ttg aag cag agg	385
Arg Val Pro Phe Gln Gly Ser Gln Gln Ser Gln Gln Leu Lys Gln Arg	
115 120 125	
gct cag att ctc cct aac gta tgc aac ctt caa tca aga cga tgc gaa	433
Ala Gln Ile Leu Pro Asn Val Cys Asn Leu Gln Ser Arg Arg Cys Glu	
130 135 140	
atc gaa agc gtc agg agt gtt gct gag agc aat ttt gaa atc cca ttt	481
Ile Glu Ser Val Arg Ser Val Ala Glu Ser Asn Phe Glu Ile Pro Phe	
145 150 155 160	
gat atg ccg ttt gat atc cct tgg ccc ttt cgc cca agc tca gag tca	529
Asp Met Pro Phe Asp Ile Pro Trp Pro Phe Arg Pro Ser Ser Glu Ser	
165 170 175	
cag caa tgc aga cag agt gaa atc caa agg cct gtg agt cag tgc caa	577
Gln Gln Cys Arg Gln Ser Glu Ile Gln Arg Pro Val Ser Gln Cys Gln	
180 185 190	
agg tat gtg gag cag caa att cag tcc tcc agg cca tac caa cag agc	625
Arg Tyr Val Glu Gln Gln Ile Gln Ser Ser Arg Pro Tyr Gln Gln Ser	
195 200 205	
ccg tac gac cgg agg caa cag agc cca tac gac cgg agg caa cag agc	673
Pro Tyr Asp Arg Arg Gln Gln Ser Pro Tyr Asp Arg Arg Gln Gln Ser	
210 215 220	
cca tat gaa cag agg caa gga cca tac gaa cag agg cca tac gaa cag	721
Pro Tyr Glu Gln Arg Gln Gly Pro Tyr Glu Gln Arg Pro Tyr Glu Gln	
225 230 235 240	
agg cca tac caa cag cga gga gga cga cag cag gag cag caa ggg ctc	769
Arg Pro Tyr Gln Gln Arg Gly Gly Arg Gln Gln Glu Gln Gln Gly Leu	
245 250 255	
cag caa tgc tgc aac gag ctc caa aac gtg agg agg gag tgt cag tgc	817
Gln Gln Cys Cys Asn Glu Leu Gln Asn Val Arg Arg Glu Cys Gln Cys	
260 265 270	
gag gcg att aag gaa gtg ggc caa aga atg agg cag cag caa caa caa	865
Glu Ala Ile Lys Glu Val Gly Gln Arg Met Arg Gln Gln Gln Gln Gln	
275 280 285	
caa cgt agg cag tat ggt ggg cag cag aca caa act gtg gag aga att	913
Gln Arg Arg Gln Tyr Gly Gly Gln Gln Thr Gln Thr Val Glu Arg Ile	
290 295 300	

59

ctt gag aat ctg cct aac caa tgc gac cta gat gtc cag caa tgc aac 961  
 Leu Glu Asn Leu Pro Asn Gln Cys Asp Leu Asp Val Gln Gln Cys Asn  
 305 310 315 320

atc ccc tac tga 973  
 Ile Pro Tyr

&lt;210&gt; 52

&lt;211&gt; 323

&lt;212&gt; PRT

&lt;213&gt; Helianthus annuus

&lt;400&gt; 52

Ala Lys Ile Thr Leu Leu Leu Leu Ala Leu Ala Ala Leu Val Ala Leu  
 1 5 10 15  
 Ala Thr Ala His Thr Thr Ile Ile Thr Thr Thr Ile Asp Asp Glu Asn  
 20 25 30  
 Pro Ile Ser Glu Gln Arg Gln Cys Trp Gln Gln Val Gln Gly Gln Arg  
 35 40 45  
 Leu Asn Gln Cys Arg Met Phe Leu Gln Gln Gly Gln Arg Gly Gln Gln  
 50 55 60  
 His Gln Gln Gln Gln His Gln Gln Gln Glu Gln Gln Leu Leu Gln Gln  
 65 70 75 80  
 Cys Cys Gln Glu Leu Gln Asn Ile Glu Gly Gln Cys Gln Cys Glu Ala  
 85 90 95  
 Val Lys Gln Val Val Arg Asp Ala Gln Arg His Glu Gln Gln Arg Pro  
 100 105 110  
 Arg Val Pro Phe Gln Gly Ser Gln Gln Ser Gln Gln Leu Lys Gln Arg  
 115 120 125  
 Ala Gln Ile Leu Pro Asn Val Cys Asn Leu Gln Ser Arg Arg Cys Glu  
 130 135 140  
 Ile Glu Ser Val Arg Ser Val Ala Glu Ser Asn Phe Glu Ile Pro Phe  
 145 150 155 160  
 Asp Met Pro Phe Asp Ile Pro Trp Pro Phe Arg Pro Ser Ser Glu Ser  
 165 170 175  
 Gln Gln Cys Arg Gln Ser Glu Ile Gln Arg Pro Val Ser Gln Cys Gln  
 180 185 190  
 Arg Tyr Val Glu Gln Gln Ile Gln Ser Ser Arg Pro Tyr Gln Gln Ser  
 195 200 205  
 Pro Tyr Asp Arg Arg Gln Gln Ser Pro Tyr Asp Arg Arg Gln Gln Ser  
 210 215 220  
 Pro Tyr Glu Gln Arg Gln Gly Pro Tyr Glu Gln Arg Pro Tyr Glu Gln  
 225 230 235 240  
 Arg Pro Tyr Gln Gln Arg Gly Gly Arg Gln Gln Glu Gln Gln Gly Leu  
 245 250 255  
 Gln Gln Cys Cys Asn Glu Leu Gln Asn Val Arg Arg Glu Cys Gln Cys  
 260 265 270  
 Glu Ala Ile Lys Glu Val Gly Gln Arg Met Arg Gln Gln Gln Gln  
 275 280 285  
 Gln Arg Arg Gln Tyr Gly Gly Gln Gln Thr Gln Thr Val Glu Arg Ile  
 290 295 300

Leu Glu Asn Leu Pro Asn Gln Cys Asp Leu Asp Val Gln Gln Cys Asn  
 305 310 315 320  
 Ile Pro Tyr

&lt;210&gt; 53

&lt;211&gt; 1114

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: DNA  
 construct coding for dsRNA

&lt;400&gt; 53

```

ggccgcgtgt tccatttggc cggaaacaac cagcagggag gctttggcgg ttcacagcaa 60
caacaagaac agaaaaactt gtggagcggg ttcgacgcac aggtcatagc tcaagcattg 120
aaaattgacg ttcagttggc tcagcagcct cagaaccaac aagacagcag aggaaacatc 180
gttcgtgtta agggacctt ccaggtcgtg aggccacctc taagacagcc ctacgagagc 240
gaggagtggg gacacccacg tagccacag ggcaacggcc ttgaggagac tatctgcagc 300
atgaggtccc acgagaacat tgacgacctt gctcgtgctg acgtgtacaa gcccagccta 360
ggtcgcgtga ccagcgtcaa cagctatacc ttgcccattc tggagtatgt caggctcagt 420
gccactcgtg gcgttctcca gggtggtacc ttctgtaaca ttgacaaaa catgtgaaca 480
cgatcatccgt catatagaac ttccaatttt aatatgtttt gctaaagaaa aaaaaaggga 540
ataaatatct atcaaatcca tttttaaaac atttggtatac gttcttaaat aatttaggat 600
atgactaatt tttctttttg gtaaaaatgt taatatctat atttaattta ttaagaaaaa 660
tgtacttaca ccctggagaa cgccacgagt ggcactgagc ctgacatact ccaagatggg 720
caaggtatag ctggttgacgc tggtcacgcg acctaggtct ggcttgtaga cgtcagcacg 780
agcagggctc tcaatgttct cgtgggacct catgctgcag atagtctcct caaggccgtt 840
gccctgtggg ctacgtgggt gtctccactc ctgcgtctcg tagggctgtc ttagaggtgg 900
cctcacgacc tggaaaggct ccttaacacg aacgatgttt cctctgctgt cttgttggtt 960
ctgaagctgc tgagccaact gaacgtcaat tttcaatgct tgagctatga cctgtgcgtc 1020
gaaccgcgtc cacaagtttt tctgttcttg ttgttgctgt gaaccgcaa agcctccctg 1080
ctggttggtt ccggccaaat ggaacacgcg gccg 1114

```

&lt;210&gt; 54

&lt;211&gt; 1114

&lt;212&gt; RNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: RNA sequence  
 for forming dsRNA

&lt;400&gt; 54

```

ggccgcgugu uccauuuggc cggaaacaac cagcagggag gcuuuggcgg uucacagcaa 60
caacaagaac agaaaaacu uggagcggg uucgacgcac aggucauagc ucaagcauug 120
aaaauugacg uucaguuggc ucagcagcuu cagaaccaac aagacagcag aggaaacauc 180
guucguguua agggaccuuu ccaggucgug aggccaccuc uaagacagcc cuacgagagc 240
gaggagugga gacacccacg uagcccacag ggcaacggcc uugaggagac uaucugcagc 300
augagguccc acgagaacau ugacgaccu gcucgugcug acguguacaa gcccagccua 360
ggucgcguga ccagcguaa cagcuauacc uugcccaucu uggagauugu caggcucagu 420
gccacucgug gcguucucca ggguggaucc uucuguaaca uuugacaaaa caugugaaca 480
cguaucuccu cauauagaac uuccaauuuu aaauuguuuu gcuaaagaaa aaaaaaggga 540
auaaauaucu aucaauuua uuuuuuuuac auuuguauac guucuuaauu aauuuaggau 600
augacuaauu uuucuauuug guaaaaaugu uaauaucuau auuuauuuu uuaagaaaaa 660
uguacuuaa cccuggagaa cgccacgagu ggcacugagc cugacauacu ccaagauggg 720
caagguauag cuguugacgc uggucacgcg accuaggcug ggcuuuguaca cgucagcacg 780
agcagggucg ucaauguucu cgugggaccu caugcugcag auagucuccu caaggccguu 840

```

## 61

```

gccuguggg cuacgugggu guccuccacuc cucgcucucg uagggcuguc uuagaggugg 900
ccucacgacc uggaaagguc ccuaaacacg aacgauguuu ccucugcugu cuuguugguu 960
cugaagcugc ugagccaacu gaacgucaau uuuccaagcu ugagcuauga ccugugcguc 1020
gaacccgcuc cacaaguuuu ucuguucuuu uuguugcugu gaaccgcca agccucccug 1080
cugguuguuu ccggccaaau ggaacacgcg gccg                                     1114

```

&lt;210&gt; 55

&lt;211&gt; 1789

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: DNA  
construct coding for dsRNA

&lt;400&gt; 55

```

gcggccgcgg atcctcaggg tcttttcttg cccactttct tgaacgccgg caaactcacg 60
tttgttggtt acggaagggg tctaattgga agagttattc cgggatgcgc cgagacgttc 120
atggagtcac cggtatattg agaaggtcaa ggtcagggtc agagtcaagg gttccgtgac 180
atgcaccaga aagtagagca cctacggtgc ggtgacacca ttgcaacacc atctggtgta 240
gctcaatggt tctacaacaa tggaaatgag cctctcattc ttgttgacgc cgcggtatctc 300
gccagcaacc agaaccagct tgaccgcaac cttagaccat ttttgatagc cggaacaac 360
ccacaagggc aggaatggct acaaggccga aagcaacaga agcaaaacaa catcttcaat 420
ggcttcgcac ctgagatctt ggctcaagcc ttcaagatca atgtcgagac ggctcagcag 480
ctccagaacc agcaagataa ccgtggcaac atcgtcaagg tcaacggacc ttccggcgtc 540
attaggccac ctttgagacg cggcgaggc ggccaacaac cacatgaaat agctaattggt 600
ttagaggaga tttgtgcac catgcgatgc actgaaaacc tcgatgacce gtcggatgct 660
gacgtgtaca agccatcact cggatacatt agcacactta acagctacaa tcttcctatc 720
ctcagacttc tccgccttag cgtctctcgt ggctccatcc gtaaaactcg aggtaagctc 780
aacaatctt tagaaaatta attttatgtg acatatgcaa taatttgatt tggcaagata 840
aactaataga ttttgcgatt tggagtttta aactctaaat aatctaaatc gttttcaatt 900
ggtttaaata tatactcttg atttttaatc gtttttaatt aaaaaatata tatatatata 960
tatactcttg atttttaatc gttttcaatt taaaaaatat cttgcacgca gaacgctgtc 1020
gagttttacg gatggagcca cgaagagcgc taaggcggag aagtctgagg ataggaagat 1080
tgtagctggt aagtgtgcta atgtatccga gtgatggctt gtacacgtca gcacccgacg 1140
ggtcacgcag gttttcagtg catcgcatgg tgcacaaagt ctctctaaa ccattagcta 1200
tttcatgtgg ttgttgggcc ccttcgccgc gtctcaaggg tggcctaatt acgccgaaag 1260
gtccgttgac cttgacgatg ttgccacggt tatcttgctg gttctggagc tgctgagccg 1320
tctcgacatt gatcttgaag gcttgagcca agatctcagg tgccaagcca ttgaagatgt 1380
tgttttgctt ctgttgcttt cggccttgta gccattcctg cccttggtggg ttgtttccgg 1440
ctatcaaaaa tgggtctaagg ttgcggtcaa gctggttctg gttgctggcg agatccgcgg 1500
ctgcaacaag aatgagaggg tcatctccat tgtttagtaa ccattgagct acaccagatg 1560
gtgttgcaat ggtgtcaccg caccgtaggt gctctacttt ctggtgcatg tcacggaacc 1620
cttgactctg accctgacct tgaccttctc caaataaccg tgactccatg aacgtctcgg 1680
cgcatcccgg aataactctt cccattagac cccttcctgt aacaacaaac gtgagtttgc 1740
cggcgttcaa gaaagtgggc aagaaaagac cctgaggatc cgcggccgc                                     1789

```

&lt;210&gt; 56

&lt;211&gt; 1789

&lt;212&gt; RNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: : RNA  
sequence for forming dsRNA

&lt;400&gt; 56

```

gcggccgcgg auccucaggg ucuuuucuuu cccacuuucu ugaacgccgg caaacucacg 60
uuuguuguuc acggaagggg ucuaauggga agaguauuuc cgggaugcgc cgagacguuc 120
auggagucac cgguaauugg agaaggucac ggucaggguc agagucaagg guuccgugac 180
augcaccaga aaguagagca ccuacggugc ggugacacca uugcaacacc aucuggugua 240

```

```

gcuccauggu ucuacaacaa uggaaaugag ccucucauuc uuguugcagc cgcgggaucuc 300
gccagcaacc agaaccagcu ugaccgcaac cuuagaccuau uuuugauagc cggaaacaac 360
ccacaagggc aggaauaggcu acaaggccga aagcaacaga agcaaaacaa caucuucuu 420
ggcuucgcac cugagauuuu ggcucaagcc uucaagauga augucgagac ggcucagcag 480
cuccagaacc agcaagauaa ccguggcaac aucgucaagg ucaacggacc uuucggcguc 540
auuaggccac ccuugagacg cggcggaagg ggccaacaac cacaugaaau agcuuauggu 600
uuagaggaga cuuugugcac caugcgauug acugaaaacc ucgaugaccc gucggaugcu 660
gacguguaca agccaucacu cggauacauu agcacacuua acagcuacaa ucuuccuau 720
cucagacuuc uccgccuuag cgcucuucgu ggcuccaucc guaaaacucg agguaaagcuc 780
aacaauucuu uagaaaauua auuuuauug acauugcaa uaauuugauu uggcaagaua 840
aacuaauaga uuuugcgauu uggaguuuua aacucuaaa auucuaaauc guuuucaauu 900
gguuuaaua uauaucuugc auuuuuaua guuuuuauu aaaaaauua uauauauua 960
uauaucuugc auuuuuaua guuuucaauu uaaaaauua cuugcacgca gaacgcuguc 1020
gaguuuuacg gauggagcca cgaagagcgc uaaggcggag aagucugagg auaggaagau 1080
uguagcuguu aagugugcua auguauccga gugauggcuu guacacguca gcauccgacg 1140
ggucaucgag guuuucagug caucgcaugg ugcacaaagu cuccucuaaa ccuuuagcua 1200
uuucaugugg uuguuggccg ccuucgcccg gucucaaggg uggccuaaug acgccgaaag 1260
guccguugac cuugacgaug uugccacggg uaucuugcug guucuggagc ugcugagccg 1320
ucucgacauu gaucuugaag gcuugagcca agaucucagg ugcgaagcca uugaagaugu 1380
uguuuugcuu cuguugcuuu cggccuugua gccauuccug ccuuguggg uuguuuccgg 1440
cuaucaaaaa uggucuaagg uugcggucua gcugguucug guugcuggcg agauccgag 1500
cugcaacaag aaugagaggc ucauuuccau uguuguagaa ccuugagcu acaccagcg 1560
guguugcaau ggugucaccg caccguaggu gcucuacuuu cuggugcaug ucacggaacc 1620
cuugacucug acccugaccu ugaccuucuc caaaauaccg ugacuccaug aacgucucgg 1680
cgcaucccgg aauaacucuu ccuauuagac ccuuccgug aacaacaaac gugaguuuugc 1740
cggcguucaa gaaagugggc aagaaaagac ccugaggauc cgcggccgc 1789

```

&lt;210&gt; 57

&lt;211&gt; 1273

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: DNA  
construct coding for dsRNA

&lt;400&gt; 57

```

gcgccgcgcg atccatggct aacaagctct tcctcgtctg cgcaactctc gccctctgct 60
tcctcctcac caacgcttcc atctaccgca cgttgtcgga attcgaagaa gatgacgcca 120
gcaaccccgat aggtccaaga cagagatgcc agaaggagtt tcagcaatca caaacaccta 180
gagcttgcca gagatggatg agcaagcaaa tgaggcaagg acgtggtggt ggtccttccc 240
tcgacgatga gttcgatctc gaggggcccc agcagggata ccagctactc cagcagtgtc 300
gcaacgagct tcgccaggaa gagccagttt gcgtttgccc caccttgaaa caagctgcca 360
gggcagtttag cctccaggga cagcacggac cattccaatc caggaaaatt taccagtcag 420
ctaagtactt gcctaacatt tgcaagatcc agcaagttgg tgaatgtccc ttccagacca 480
ccatcccttt ctccctcctt tactactagg gtactcgagg taagctcaac aaatctttag 540
aaaattaatt ttatgtgaca tatgcaataa tttgatttgg caagataaac taatagattt 600
tgcgatttgg agtttttaaac tctaaataat ctaaatcggt ttcaattggt ttaaataatat 660
atcttgcatt tttaatcggt ttttaattaaa aaatatatat atatatatat atcttgcatt 720
tttaatcggt ttcaatttaa aaaatatctt gcacgcagaa cgctgtcgac taccctagta 780
gtaaggaggg aagaaaggga tgggtggtctg gaagggacat tcaccaactt gctggatctt 840
gcaaatgtta ggcaagtact tagctgactg gtaaattttc ctggattgga atggtccgtg 900
ctgtccctgg aggctaactg ccttgccagc ttgtttcaag gtggggcaaa cgcaaaactgg 960
ctcttcctgg cgaagctcgt tgcagcactg ctggagtagc tggatatcct gctggggggc 1020
ctcgaatcgt aactcatcgt cgagggaagg accaccacca cgtccttgcc tcatttgcct 1080
gctcatccat ctctggcaag ctcttaggtg ttgtgattgc tgaaactcct tctggcatct 1140
ctgtcttggg cctacgggggt tgctggcgct atcttcttcg aattcgacaa cgggtgcggt 1200
gatggaagcg ttggtgagga ggaagcagag ggcgagagtt gcgcagacga ggaagagctt 1260
gttagccatg gat

```

1273

&lt;210&gt; 58

&lt;211&gt; 1273

&lt;212&gt; RNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: : RNA  
sequence for forming dsRNA

&lt;400&gt; 58

```

gcggccgcgg auccauggcu aacaagcucu uccucgucug cgcaacucuc gccucugcu 60
uccuccucac caacgcuucc aucuaccgca ccguugucga auucgaagaa gaugacgcca 120
gcaaccccggu agguccaaga cagagaugcc agaaggaguu ucagcaauca caacaccuaa 180
gagcuugcca gagauggaug agcaagcaaa ugaggcaagg acgugguggu gguccuuccc 240
ucgacgaugâ guucgauuuc gagggccccc agcagggaua ccagcuacuc cagcagugcu 300
gcaacgagcu ucgccaggaa gagccaguuu gcguuugccc caccuugaaa caagcugcca 360
gggcaguuag ccuccaggga cagcacggac cauuccaau caggaaaauu uaccagucag 420
cuaaguacuu gccuaacauu ugcaagauc agcaaguugg ugaauugucc uuccagacca 480
ccaucccuuu cuuccuccu uacuacuagg guacucgagg uaagcucaac aaauuuuag 540
aaaauuaauu uuaugugaca uaugcaaua uuugauuugg caagauaaac uaauagauuu 600
ugcgauuugg aguuuuuuaac ucuaauuaa cuaaaucguu uucaauuggu uuaauauau 660
aucuugcauu uuuaauucguu uuuaauuaaa aaauauauau auauauauau aucuugcauu 720
uuuaauucguu uucaauuuuaa aaaauaucuu gcacgcagaa cgcugucgac uaccuagua 780
guaaggaggga aagaaaggga ugguggucug gaagggacau ucaccaacuu gcuggaucuu 840
gcaaauguua ggcaaguacu uagcugacug guaaauuuuc cuggauugga augguccgug 900
cugucccugg aggcuaacug ccucggcagc uuguuucaag gugggggcaa cgaacucgg 960
cucuuccugg cgaagcucgu ugcagcacug cuggaguagc ugguaucuccu gcuggggggcc 1020
cucgaaaucg aacucaucgu cgagggaagg accaccacca cguccuugcc ucauuugcuu 1080
gcucauccau cucuggcaag cucuuaggug uugugauugc ugaaacuccu ucuggcaucu 1140
cugucuugga ccuacggggu ugcuggcguc aucuucucg aaucgacaa cggugcggu 1200
gauggaagcg uugugagga ggaagcagag ggcgagaguu ggcgagacga ggaagagcuu 1260
guuagccaug gau

```

&lt;210&gt; 59

&lt;211&gt; 1575

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1572)

&lt;223&gt; 12S cruciferin

&lt;400&gt; 59

```

atg gtt aag ctc agc aat ctc ctc gtt gca acc ttc ggg gtt ctc ctc 48
Met Val Lys Leu Ser Asn Leu Leu Val Ala Thr Phe Gly Val Leu Leu
  1             5             10             15

gtc ctt aac ggc tgc ctt gcg agg cag tca ctt ggg gtt cct cct cag 96
Val Leu Asn Gly Cys Leu Ala Arg Gln Ser Leu Gly Val Pro Pro Gln
          20             25             30

cta cag aac gag tgt aac ctc gac aac cta gat gtt ctc caa gcc acc 144
Leu Gln Asn Glu Cys Asn Leu Asp Asn Leu Asp Val Leu Gln Ala Thr
          35             40             45

gaa act atc aag agt gaa gcc ggt cag atc gag tac tgg gac cac aac 192
Glu Thr Ile Lys Ser Glu Ala Gly Gln Ile Glu Tyr Trp Asp His Asn
          50             55             60

cac cct cag ctc cga tgt gtt ggt gtt tcc gtt gct cgt tat gta att 240
His Pro Gln Leu Arg Cys Val Gly Val Ser Val Ala Arg Tyr Val Ile
          65             70             75             80

```

64

gaa	caa	ggc	ggt	ctt	tac	ttg	ccc	acc	ttc	ttc	act	tcc	cca	aaa	att	288
Glu	Gln	Gly	Gly	Leu	Tyr	Leu	Pro	Thr	Phe	Phe	Thr	Ser	Pro	Lys	Ile	
				85					90					95		
tcc	tac	gtc	gtt	caa	gga	acg	ggt	atc	agc	gga	aga	gtg	gtc	cct	gga	336
Ser	Tyr	Val	Val	Gln	Gly	Thr	Gly	Ile	Ser	Gly	Arg	Val	Val	Pro	Gly	
			100					105					110			
tgt	gcc	gag	acc	ttc	atg	gac	tcg	cag	ccg	atg	caa	gga	caa	caa	caa	384
Cys	Ala	Glu	Thr	Phe	Met	Asp	Ser	Gln	Pro	Met	Gln	Gly	Gln	Gln	Gln	
		115					120					125				
ggc	caa	cca	tgg	caa	gga	cga	cag	gga	caa	caa	ggc	caa	cca	tgg	gaa	432
Gly	Gln	Pro	Trp	Gln	Gly	Arg	Gln	Gly	Gln	Gln	Gly	Gln	Pro	Trp	Glu	
	130					135					140					
gga	cag	gga	caa	cag	gga	caa	caa	gga	aga	caa	ggc	caa	cca	tgg	gaa	480
Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Arg	Gln	Gly	Gln	Pro	Trp	Glu	
145					150				155						160	
gga	cag	gga	caa	cag	gga	caa	caa	gga	cga	cag	gga	caa	caa	ggc	caa	528
Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Arg	Gln	Gly	Gln	Gln	Gly	Gln	
			165					170						175		
cca	tgg	gaa	gga	cag	gga	cag	cag	gga	caa	caa	ggg	ttc	cgt	gac	atg	576
Pro	Trp	Glu	Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Phe	Arg	Asp	Met	
			180					185					190			
cac	cag	aag	gtg	gaa	cat	gtg	aga	cgc	gga	gac	gtc	ttt	gcc	aac	act	624
His	Gln	Lys	Val	Glu	His	Val	Arg	Arg	Gly	Asp	Val	Phe	Ala	Asn	Thr	
		195					200					205				
cca	ggc	tct	gcc	cac	tgg	atc	tac	aac	tca	gga	gaa	cag	cca	ctt	gtc	672
Pro	Gly	Ser	Ala	His	Trp	Ile	Tyr	Asn	Ser	Gly	Glu	Gln	Pro	Leu	Val	
	210					215					220					
atc	atc	gct	ctt	ctc	gac	atc	gcc	aac	tac	caa	aac	caa	ctc	gac	cgc	720
Ile	Ile	Ala	Leu	Leu	Asp	Ile	Ala	Asn	Tyr	Gln	Asn	Gln	Leu	Asp	Arg	
225					230					235					240	
aac	cct	aga	gtg	ttc	cat	ttg	gcc	gga	aac	aac	cag	cag	gga	ggc	ttt	768
Asn	Pro	Arg	Val	Phe	His	Leu	Ala	Gly	Asn	Asn	Gln	Gln	Gly	Gly	Phe	
				245				250						255		
ggc	ggt	tca	cag	caa	caa	caa	gaa	cag	aaa	aac	ttg	tgg	agc	ggg	ttc	816
Gly	Gly	Ser	Gln	Gln	Gln	Gln	Glu	Gln	Lys	Asn	Leu	Trp	Ser	Gly	Phe	
			260					265					270			
gac	gca	cag	gtc	ata	gct	caa	gca	ttg	aaa	att	gac	gtt	cag	ttg	gct	864
Asp	Ala	Gln	Val	Ile	Ala	Gln	Ala	Leu	Lys	Ile	Asp	Val	Gln	Leu	Ala	
			275				280					285				
cag	cag	ctt	cag	aac	caa	caa	gac	agc	aga	gga	aac	atc	gtt	cgt	gtt	912
Gln	Gln	Leu	Gln	Asn	Gln	Gln	Asp	Ser	Arg	Gly	Asn	Ile	Val	Arg	Val	
		290				295					300					
aag	gga	cct	ttc	cag	gtc	gtg	agg	cca	cct	cta	aga	cag	ccc	tac	gag	960
Lys	Gly	Pro	Phe	Gln	Val	Val	Arg	Pro	Pro	Leu	Arg	Gln	Pro	Tyr	Glu	
305					310					315					320	
agc	gag	gag	tgg	aga	cac	cca	cgt	agc	cca	cag	ggc	aac	ggc	ctt	gag	1008
Ser	Glu	Glu	Trp	Arg	His	Pro	Arg	Ser	Pro	Gln	Gly	Asn	Gly	Leu	Glu	
				325				330						335		
gag	act	atc	tgc	agc	atg	agg	tcc	cac	gag	aac	att	gac	gac	cct	gct	1056
Glu	Thr	Ile	Cys	Ser	Met	Arg	Ser	His	Glu	Asn	Ile	Asp	Asp	Pro	Ala	
			340					345					350			

## 65

```

cgt gct gac gtg tac aag ccc agc cta ggt cgc gtg acc agc gtc aac      1104
Arg Ala Asp Val Tyr Lys Pro Ser Leu Gly Arg Val Thr Ser Val Asn
      355                                360                                365

agc tat acc ttg ccc atc ttg gag tat gtc agg ctc agt gcc act cgt      1152
Ser Tyr Thr Leu Pro Ile Leu Glu Tyr Val Arg Leu Ser Ala Thr Arg
      370                                375                                380

ggc gtt ctc cag ggt aat gcg atg gtg ctt cct aaa tac aac atg aac      1200
Gly Val Leu Gln Gly Asn Ala Met Val Leu Pro Lys Tyr Asn Met Asn
      385                                390                                395                                400

gct aac gag atc ttg tac tgc act gga gga caa gga agg atc caa gtg      1248
Ala Asn Glu Ile Leu Tyr Cys Thr Gly Gly Gln Gly Arg Ile Gln Val
      405                                410                                415

gtc aac gac aac gga cag aac gtg ttg gac caa cag gtg cag aag gga      1296
Val Asn Asp Asn Gly Gln Asn Val Leu Asp Gln Gln Val Gln Lys Gly
      420                                425                                430

cag ctc gtg gtc atc cca caa ggg ttc gca tac gtt gtc cag tcc cac      1344
Gln Leu Val Val Ile Pro Gln Gly Phe Ala Tyr Val Val Gln Ser His
      435                                440                                445

gga aac aag ttc gag tgg atc tct ttc aaa act aat gaa aac gca atg      1392
Gly Asn Lys Phe Glu Trp Ile Ser Phe Lys Thr Asn Glu Asn Ala Met
      450                                455                                460

atc agc act ttg gcg ggt aga acc tcg ctc ttg agg gca ttg cca ttg      1440
Ile Ser Thr Leu Ala Gly Arg Thr Ser Leu Leu Arg Ala Leu Pro Leu
      465                                470                                475                                480

gag gtc ata tca aat ggt ttc cag atc tct ccc gag gaa gct agg aag      1488
Glu Val Ile Ser Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Lys
      485                                490                                495

atc aag ttc aac aca ctt gag acc act ttg acc cgc gct gcc ggt agg      1536
Ile Lys Phe Asn Thr Leu Glu Thr Thr Leu Thr Arg Ala Ala Gly Arg
      500                                505                                510

caa caa caa cag ttg atc gag gag att gtc gag gct taa
Gln Gln Gln Gln Leu Ile Glu Glu Ile Val Glu Ala
      515                                520

```

&lt;210&gt; 60

&lt;211&gt; 524

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 60

```

Met Val Lys Leu Ser Asn Leu Leu Val Ala Thr Phe Gly Val Leu Leu
  1              5              10              15

Val Leu Asn Gly Cys Leu Ala Arg Gln Ser Leu Gly Val Pro Gln
      20              25              30

Leu Gln Asn Glu Cys Asn Leu Asp Asn Leu Asp Val Leu Gln Ala Thr
      35              40              45

Glu Thr Ile Lys Ser Glu Ala Gly Gln Ile Glu Tyr Trp Asp His Asn
      50              55              60

His Pro Gln Leu Arg Cys Val Gly Val Ser Val Ala Arg Tyr Val Ile
      65              70              75              80

Glu Gln Gly Gly Leu Tyr Leu Pro Thr Phe Phe Thr Ser Pro Lys Ile
      85              90              95

```



## 66

Ser	Tyr	Val	Val	Gln	Gly	Thr	Gly	Ile	Ser	Gly	Arg	Val	Val	Pro	Gly	100	105	110
Cys	Ala	Glu	Thr	Phe	Met	Asp	Ser	Gln	Pro	Met	Gln	Gly	Gln	Gln	Gln	115	120	125
Gly	Gln	Pro	Trp	Gln	Gly	Arg	Gln	Gly	Gln	Gln	Gly	Gln	Pro	Trp	Glu	130	135	140
Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Arg	Gln	Gly	Gln	Pro	Trp	Glu	145	150	155
Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Arg	Gln	Gly	Gln	Gln	Gly	Gln	165	170	175
Pro	Trp	Glu	Gly	Gln	Gly	Gln	Gln	Gly	Gln	Gln	Gly	Phe	Arg	Asp	Met	180	185	190
His	Gln	Lys	Val	Glu	His	Val	Arg	Arg	Gly	Asp	Val	Phe	Ala	Asn	Thr	195	200	205
Pro	Gly	Ser	Ala	His	Trp	Ile	Tyr	Asn	Ser	Gly	Glu	Gln	Pro	Leu	Val	210	215	220
Ile	Ile	Ala	Leu	Leu	Asp	Ile	Ala	Asn	Tyr	Gln	Asn	Gln	Leu	Asp	Arg	225	230	235
Asn	Pro	Arg	Val	Phe	His	Leu	Ala	Gly	Asn	Asn	Gln	Gln	Gly	Gly	Phe	245	250	255
Gly	Gly	Ser	Gln	Gln	Gln	Gln	Glu	Gln	Lys	Asn	Leu	Trp	Ser	Gly	Phe	260	265	270
Asp	Ala	Gln	Val	Ile	Ala	Gln	Ala	Leu	Lys	Ile	Asp	Val	Gln	Leu	Ala	275	280	285
Gln	Gln	Leu	Gln	Asn	Gln	Gln	Asp	Ser	Arg	Gly	Asn	Ile	Val	Arg	Val	290	295	300
Lys	Gly	Pro	Phe	Gln	Val	Val	Arg	Pro	Pro	Leu	Arg	Gln	Pro	Tyr	Glu	305	310	315
Ser	Glu	Glu	Trp	Arg	His	Pro	Arg	Ser	Pro	Gln	Gly	Asn	Gly	Leu	Glu	325	330	335
Glu	Thr	Ile	Cys	Ser	Met	Arg	Ser	His	Glu	Asn	Ile	Asp	Asp	Pro	Ala	340	345	350
Arg	Ala	Asp	Val	Tyr	Lys	Pro	Ser	Leu	Gly	Arg	Val	Thr	Ser	Val	Asn	355	360	365
Ser	Tyr	Thr	Leu	Pro	Ile	Leu	Glu	Tyr	Val	Arg	Leu	Ser	Ala	Thr	Arg	370	375	380
Gly	Val	Leu	Gln	Gly	Asn	Ala	Met	Val	Leu	Pro	Lys	Tyr	Asn	Met	Asn	385	390	395
Ala	Asn	Glu	Ile	Leu	Tyr	Cys	Thr	Gly	Gly	Gln	Gly	Arg	Ile	Gln	Val	405	410	415
Val	Asn	Asp	Asn	Gly	Gln	Asn	Val	Leu	Asp	Gln	Gln	Val	Gln	Lys	Gly	420	425	430
Gln	Leu	Val	Val	Ile	Pro	Gln	Gly	Phe	Ala	Tyr	Val	Val	Gln	Ser	His	435	440	445
Gly	Asn	Lys	Phe	Glu	Trp	Ile	Ser	Phe	Lys	Thr	Asn	Glu	Asn	Ala	Met	450	455	460
Ile	Ser	Thr	Leu	Ala	Gly	Arg	Thr	Ser	Leu	Leu	Arg	Ala	Leu	Pro	Leu	465	470	475
																		480

67

```
<210> 61
<211> 1419
<212> DNA
<213> Arabidopsis thaliana
<220>
<221> CDS
<222> (1)..(1416)
<223> 12S Cra1 storage protein
<400> 61
```

<400> 61																
atg gct cga gtc tct tct ctt ctt tct ttc tgc tta aca ctt ttg atc	48															
Met Ala Arg Val Ser Ser Leu Leu Ser Phe Cys Leu Thr Leu Leu Ile																
1 5 10 15																
ctt ttc cat ggc tac gcg gct caa cag ggt cag cag ggt cag cag ttt	96															
Leu Phe His Gly Tyr Ala Ala Gln Gln Gly Gln Gln Gly Gln Gln Phe																
20 25 30																
ccg aac gag tgc cag ctc gac cag ctc aat gcg ctc gag ccg tca cac	144															
Pro Asn Glu Cys Gln Leu Asp Gln Leu Asn Ala Leu Glu Pro Ser His																
35 40 45																
gta ctg aag agc gag gct ggt cgc atc gag gtg tgg gac cac cac gct	192															
Val Leu Lys Ser Glu Ala Gly Arg Ile Glu Val Trp Asp His His Ala																
50 55 60																
cct cag ctc cgt tgc tca ggt gtc tcc ttt gca cgt tac atc atc gag	240															
Pro Gln Leu Arg Cys Ser Gly Val Ser Phe Ala Arg Tyr Ile Ile Glu																
65 70 75 80																
tct aag ggt ctc tac ttg ccc tct ttc ttt aac acc gcg aag ctc tct	288															
Ser Lys Gly Leu Tyr Leu Pro Ser Phe Phe Asn Thr Ala Lys Leu Ser																
85 90 95																
ttc gtg gct aag gga cga ggt ctt atg gga aaa gtg atc cct gga tgc	336															
Phe Val Ala Lys Gly Arg Gly Leu Met Gly Lys Val Ile Pro Gly Cys																
100 105 110																
gcc gaa aca ttc caa gac tca tca gag ttc caa cca cgc ttc gaa ggt	384															
Ala Glu Thr Phe Gln Asp Ser Ser Glu Phe Gln Pro Arg Phe Glu Gly																
115 120 125																
caa ggt caa agc cag agg ttc cgt gac atg cac cag aaa gtg gag cac	432															
Gln Gly Gln Ser Gln Arg Phe Arg Asp Met His Gln Lys Val Glu His																
130 135 140																
att agg agc ggt gat acc att gcc aca aca ccc ggt gta gca cag tgg	480															
Ile Arg Ser Gly Asp Thr Ile Ala Thr Thr Pro Gly Val Ala Gln Trp																
145 150 155 160																
ttc tac aac gac gga cag cag cca ctt gtc atc gtc agc gtc ttc gat	528															
Phe Tyr Asn Asp Gly Gln Gln Pro Leu Val Ile Val Ser Val Phe Asp																
165 170 175																
cta gcc agt cac cag aac cag ctt gac cgc aac cca agg cca ttt tac	576															
Leu Ala Ser His Gln Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr																
180 185 190																

tta gcc gga aac aac cca caa ggt caa gta tgg cta caa gga cga gag	624
Leu Ala Gly Asn Asn Pro Gln Gly Gln Val Trp Leu Gln Gly Arg Glu	
195 200 205	
caa cag cca cag aag aac att ttc aat gga ttt gga ccc gag gtt att	672
Gln Gln Pro Gln Lys Asn Ile Phe Asn Gly Phe Gly Pro Glu Val Ile	
210 215 220	
gct caa gct ttg aag atc gat ctt cag aca gca cag caa ctt cag aac	720
Ala Gln Ala Leu Lys Ile Asp Leu Gln Thr Ala Gln Gln Leu Gln Asn	
225 230 235 240	
caa gat gac aac cgt gga aac att gtc cga gtc caa gga ccg ttc ggt	768
Gln Asp Asp Asn Arg Gly Asn Ile Val Arg Val Gln Gly Pro Phe Gly	
245 250 255	
gtc att agg ccg cct ttg agg ggc cag aga cct cag gag gag gaa gaa	816
Val Ile Arg Pro Pro Leu Arg Gly Gln Arg Pro Gln Glu Glu Glu Glu	
260 265 270	
gaa gaa gga cga cat gga cga cac ggt aat ggc tta gag gag acc atc	864
Glu Glu Gly Arg His Gly Arg His Gly Asn Gly Leu Glu Glu Thr Ile	
275 280 285	
tgc agc gcc agg tgc acc gat aac ctc gat gac ccg tct cgt gct gac	912
Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Arg Ala Asp	
290 295 300	
gtg tac aag cca cag ctc ggt tac atc agc act ctc aac agt tac gat	960
Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp	
305 310 315 320	
ctc ccc att ctt cgc ttc atc cgt ctc tca gcc ctc cgt gga tct atc	1008
Leu Pro Ile Leu Arg Phe Ile Arg Leu Ser Ala Leu Arg Gly Ser Ile	
325 330 335	
cgt caa aac gca atg gtg ctt cca cag tgg aac gca aac gcg aac gct	1056
Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala	
340 345 350	
att ctt tac gag aca gac ggg gaa gcc caa atc cag atc gta aac gac	1104
Ile Leu Tyr Glu Thr Asp Gly Glu Ala Gln Ile Gln Ile Val Asn Asp	
355 360 365	
aat ggt aac aga gtg ttt gac gga caa gtc tct caa gga cag ctc ata	1152
Asn Gly Asn Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Ile	
370 375 380	
gcc gta cca caa ggt ttc tcg gtg gtg aaa cgc gca aca agc aac cga	1200
Ala Val Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Asn Arg	
385 390 395 400	
ttc cag tgg gtt gag ttc aaa aca aac gct aac gcg caa atc aac act	1248
Phe Gln Trp Val Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr	
405 410 415	
ctg gcg gga cga acc tca gtc ttg aga ggt tta cca ctt gaa gtc ata	1296
Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile	
420 425 430	
acc aat ggg ttc caa atc tca ccc gaa gaa gca agg agg gtc aag ttc	1344
Thr Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Arg Val Lys Phe	
435 440 445	
aac acg ctc gag acc act ttg act cac agc agt ggc cca gct agc tac	1392
Asn Thr Leu Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr	
450 455 460	

<400>	62														
Met 1	Ala	Arg	Val	Ser 5	Ser	Leu	Leu	Ser	Phe 10	Cys	Leu	Thr	Leu	Leu	Ile 15
Leu	Phe	His	Gly 20	Tyr	Ala	Ala	Gln	Gln	Gly 25	Gln	Gln	Gly	Gln	Gln	Phe
Pro	Asn	Glu 35	Cys	Gln	Leu	Asp	Gln 40	Leu	Asn	Ala	Leu	Glu 45	Pro	Ser	His
Val 50	Leu	Lys	Ser	Glu	Ala	Gly 55	Arg	Ile	Glu	Val	Trp 60	Asp	His	His	Ala
Pro 65	Gln	Leu	Arg	Cys	Ser 70	Gly	Val	Ser	Phe	Ala 75	Arg	Tyr	Ile	Ile	Glu 80
Ser	Lys	Gly	Leu	Tyr 85	Leu	Pro	Ser	Phe	Phe 90	Asn	Thr	Ala	Lys	Leu	Ser
Phe	Val	Ala	Lys 100	Gly	Arg	Gly	Leu	Met 105	Gly	Lys	Val	Ile	Pro 110	Gly	Cys
Ala	Glu	Thr 115	Phe	Gln	Asp	Ser	Ser 120	Glu	Phe	Gln	Pro	Arg 125	Phe	Glu	Gly
Gln 130	Gly	Gln	Ser	Gln	Arg	Phe 135	Arg	Asp	Met	His	Gln 140	Lys	Val	Glu	His
Ile 145	Arg	Ser	Gly	Asp	Thr 150	Ile	Ala	Thr	Thr	Pro 155	Gly	Val	Ala	Gln	Trp 160
Phe	Tyr	Asn	Asp 165	Gly	Gln	Gln	Pro	Leu	Val 170	Ile	Val	Ser	Val	Phe	Asp 175
Leu	Ala	Ser 180	His	Gln	Asn	Gln	Leu	Asp 185	Arg	Asn	Pro	Arg 190	Pro	Phe	Tyr
Leu	Ala	Gly 195	Asn	Asn	Pro	Gln	Gly 200	Gln	Val	Trp	Leu	Gln 205	Gly	Arg	Glu
Gln 210	Gln	Pro	Gln	Lys	Asn	Ile 215	Phe	Asn	Gly	Phe	Gly 220	Pro	Glu	Val	Ile
Ala 225	Gln	Ala	Leu	Lys	Ile 230	Asp	Leu	Gln	Thr	Ala 235	Gln	Gln	Leu	Gln	Asn 240
Gln	Asp	Asp	Asn	Arg 245	Gly	Asn	Ile	Val	Arg 250	Val	Gln	Gly	Pro	Phe	Gly 255
Val	Ile	Arg 260	Pro	Pro	Leu	Arg	Gly 265	Gln	Arg	Pro	Gln	Glu 270	Glu	Glu	Glu
Glu	Glu	Gly 275	Arg	His	Gly	Arg	His 280	Gly	Asn	Gly	Leu	Glu 285	Glu	Thr	Ile
Cys 290	Ser	Ala	Arg	Cys	Thr	Asp 295	Asn	Leu	Asp	Asp	Pro 300	Ser	Arg	Ala	Asp
Val 305	Tyr	Lys	Pro	Gln	Leu	Gly 310	Tyr	Ile	Ser	Thr 315	Leu	Asn	Ser	Tyr	Asp 320

## 70

Leu Pro Ile Leu Arg Phe Ile Arg Leu Ser Ala Leu Arg Gly Ser Ile  
                                   325                                  330                                  335  
 Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala  
                                   340                                  345                                  350  
 Ile Leu Tyr Glu Thr Asp Gly Glu Ala Gln Ile Gln Ile Val Asn Asp  
                                   355                                  360                                  365  
 Asn Gly Asn Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Ile  
                                   370                                  375                                  380  
 Ala Val Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Asn Arg  
 385                                  390                                  395                                  400  
 Phe Gln Trp Val Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr  
                                   405                                  410                                  415  
 Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile  
                                   420                                  425                                  430  
 Thr Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Arg Val Lys Phe  
                                   435                                  440                                  445  
 Asn Thr Leu Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr  
                                   450                                  455                                  460  
 Gly Arg Pro Arg Val Ala Ala Ala  
 465                                  470

&lt;210&gt; 63

&lt;211&gt; 1419

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1416)

&lt;223&gt; At5g442120/MLN1\_4 storage protein

&lt;400&gt; 63

atg gct cga gtc tct tct ctt ctt tct ttc tgc tta aca ctt ttg atc	48
Met Ala Arg Val Ser Ser Leu Leu Ser Phe Cys Leu Thr Leu Leu Ile	
1                                  5                                  10                                  15	
ctt ttc cat ggc tac gcg gct caa cag ggt cag cag ggt cag cag ttt	96
Leu Phe His Gly Tyr Ala Ala Gln Gln Gly Gln Gln Gly Gln Gln Phe	
20                                  25                                  30	
ccg aac gag tgc cag ctc gac cag ctc aat gcg ctc gag ccg tca cac	144
Pro Asn Glu Cys Gln Leu Asp Gln Leu Asn Ala Leu Glu Pro Ser His	
35                                  40                                  45	
gta ctg aag agc gag gct ggt cgc atc gag gtg tgg gac cac cac gct	192
Val Leu Lys Ser Glu Ala Gly Arg Ile Glu Val Trp Asp His His Ala	
50                                  55                                  60	
cct cag ctc cgt tgc tca ggt gtc tcc ttt gca cgt tac atc atc gag	240
Pro Gln Leu Arg Cys Ser Gly Val Ser Phe Ala Arg Tyr Ile Ile Glu	
65                                  70                                  75                                  80	
tct aag ggt ctc tac ttg ccc tct ttc ttt aac acc gcg aag ctc tct	288
Ser Lys Gly Leu Tyr Leu Pro Ser Phe Phe Asn Thr Ala Lys Leu Ser	
85                                  90                                  95	
ttc gtg gct aag gga cga ggt ctt atg gga aaa gtg atc cct gga tgc	336
Phe Val Ala Lys Gly Arg Gly Leu Met Gly Lys Val Ile Pro Gly Cys	
100                                  105                                  110	

gcc gaa aca ttc caa gac tca tca gag ttc caa cca cgc ttc gaa ggt	384
Ala Glu Thr Phe Gln Asp Ser Ser Glu Phe Gln Pro Arg Phe Glu Gly	
115 120 125	
caa ggt caa agc cag agg ttc cgt gac atg cac cag aaa gtg gag cac	432
Gln Gly Gln Ser Gln Arg Phe Arg Asp Met His Gln Lys Val Glu His	
130 135 140	
att agg agc ggt gat acc att gcc aca aca ccc ggt gta gca cag tgg	480
Ile Arg Ser Gly Asp Thr Ile Ala Thr Thr Pro Gly Val Ala Gln Trp	
145 150 155 160	
ttc tac aac gac gga cag gaa cca ctt gtc atc gtc agc gtc ttc gat	528
Phe Tyr Asn Asp Gly Gln Glu Pro Leu Val Ile Val Ser Val Phe Asp	
165 170 175	
cta gcc agt cac cag aac cag ctt gac cgc aac cca agg cca ttt tac	576
Leu Ala Ser His Gln Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr	
180 185 190	
tta gcc gga aac aac cca caa ggt caa gta tgg cta caa gga cga gag	624
Leu Ala Gly Asn Asn Pro Gln Gly Gln Val Trp Leu Gln Gly Arg Glu	
195 200 205	
caa cag cca cag aag aac att ttc aat gga ttt gga ccc gag gtt att	672
Gln Gln Pro Gln Lys Asn Ile Phe Asn Gly Phe Gly Pro Glu Val Ile	
210 215 220	
gct caa gct ttg aag atc gat ctt cag aca gca cag caa ctt cag aac	720
Ala Gln Ala Leu Lys Ile Asp Leu Gln Thr Ala Gln Gln Leu Gln Asn	
225 230 235 240	
caa gat gac aac cgt gga aac att gtc cga gtc caa gga ccg ttc ggt	768
Gln Asp Asp Asn Arg Gly Asn Ile Val Arg Val Gln Gly Pro Phe Gly	
245 250 255	
gtc att agg ccg cct ttg agg ggc cag aga cct cag gag gag gaa gaa	816
Val Ile Arg Pro Pro Leu Arg Gly Gln Arg Pro Gln Glu Glu Glu Glu	
260 265 270	
gaa gaa gga cga cat gga cga cac ggt aat ggc tta gag gag acc atc	864
Glu Glu Gly Arg His Gly Arg His Gly Asn Gly Leu Glu Glu Thr Ile	
275 280 285	
tgc agc gcc agg tgc acc gat aac ctc gat gac ccg tct cgt gct gac	912
Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Arg Ala Asp	
290 295 300	
gtg tac aag cca cag ctc ggt tac atc agc act ctc aac agt tac gat	960
Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp	
305 310 315 320	
ctc ccc atc ctt cgc ttc atc cgt ctc tca gcc ctc cgt gga tct atc	1008
Leu Pro Ile Leu Arg Phe Ile Arg Leu Ser Ala Leu Arg Gly Ser Ile	
325 330 335	
cgt caa aac gca atg gtg ctt cca cag tgg aac gca aac gcg aac gct	1056
Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala	
340 345 350	
att ctt tac gtg aca gac ggg gaa gcc caa atc cag atc gta aac gac	1104
Ile Leu Tyr Val Thr Asp Gly Glu Ala Gln Ile Gln Ile Val Asn Asp	
355 360 365	
aat ggt aac aga gtg ttt gac gga caa gtc tct caa gga cag ctc ata	1152
Asn Gly Asn Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Ile	
370 375 380	

72

gcc gta cca caa ggt ttc tcg gtg gtg aaa cgc gca aca agc aac cga 1200  
 Ala Val Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Asn Arg  
 385 390 395 400  
 ttc cag tgg gtt gag ttc aaa aca aac gct aac gcg caa atc aac act 1248  
 Phe Gln Trp Val Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr  
 405 410 415  
 ctg gcg gga cga acc tca gtc ttg aga ggt tta cca ctt gaa gtc ata 1296  
 Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile  
 420 425 430  
 acc aat ggg ttc caa atc tca ccc gaa gaa gca agg agg gtc aag ttc 1344  
 Thr Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Arg Val Lys Phe  
 435 440 445  
 aac acg ctc gag acc act ttg act cac agc agt ggc cca gct agc tac 1392  
 Asn Thr Leu Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr  
 450 455 460  
 gga agg cca agg gtg gct gca gct taa 1419  
 Gly Arg Pro Arg Val Ala Ala Ala  
 465 470  
 <210> 64  
 <211> 472  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 64  
 Met Ala Arg Val Ser Ser Leu Leu Ser Phe Cys Leu Thr Leu Leu Ile  
 1 5 10 15  
 Leu Phe His Gly Tyr Ala Ala Gln Gln Gly Gln Gln Gly Gln Gln Phe  
 20 25 30  
 Pro Asn Glu Cys Gln Leu Asp Gln Leu Asn Ala Leu Glu Pro Ser His  
 35 40 45  
 Val Leu Lys Ser Glu Ala Gly Arg Ile Glu Val Trp Asp His His Ala  
 50 55 60  
 Pro Gln Leu Arg Cys Ser Gly Val Ser Phe Ala Arg Tyr Ile Ile Glu  
 65 70 75 80  
 Ser Lys Gly Leu Tyr Leu Pro Ser Phe Phe Asn Thr Ala Lys Leu Ser  
 85 90 95  
 Phe Val Ala Lys Gly Arg Gly Leu Met Gly Lys Val Ile Pro Gly Cys  
 100 105 110  
 Ala Glu Thr Phe Gln Asp Ser Ser Glu Phe Gln Pro Arg Phe Glu Gly  
 115 120 125  
 Gln Gly Gln Ser Gln Arg Phe Arg Asp Met His Gln Lys Val Glu His  
 130 135 140  
 Ile Arg Ser Gly Asp Thr Ile Ala Thr Thr Pro Gly Val Ala Gln Trp  
 145 150 155 160  
 Phe Tyr Asn Asp Gly Gln Glu Pro Leu Val Ile Val Ser Val Phe Asp  
 165 170 175  
 Leu Ala Ser His Gln Asn Gln Leu Asp Arg Asn Pro Arg Pro Phe Tyr  
 180 185 190  
 Leu Ala Gly Asn Asn Pro Gln Gly Gln Val Trp Leu Gln Gly Arg Glu  
 195 200 205

## 73

Gln Gln Pro Gln Lys Asn Ile Phe Asn Gly Phe Gly Pro Glu Val Ile  
 210 215 220  
 Ala Gln Ala Leu Lys Ile Asp Leu Gln Thr Ala Gln Gln Leu Gln Asn  
 225 230 235 240  
 Gln Asp Asp Asn Arg Gly Asn Ile Val Arg Val Gln Gly Pro Phe Gly  
 245 250 255  
 Val Ile Arg Pro Pro Leu Arg Gly Gln Arg Pro Gln Glu Glu Glu Glu  
 260 265 270  
 Glu Glu Gly Arg His Gly Arg His Gly Asn Gly Leu Glu Glu Thr Ile  
 275 280 285  
 Cys Ser Ala Arg Cys Thr Asp Asn Leu Asp Asp Pro Ser Arg Ala Asp  
 290 295 300  
 Val Tyr Lys Pro Gln Leu Gly Tyr Ile Ser Thr Leu Asn Ser Tyr Asp  
 305 310 315 320  
 Leu Pro Ile Leu Arg Phe Ile Arg Leu Ser Ala Leu Arg Gly Ser Ile  
 325 330 335  
 Arg Gln Asn Ala Met Val Leu Pro Gln Trp Asn Ala Asn Ala Asn Ala  
 340 345 350  
 Ile Leu Tyr Val Thr Asp Gly Glu Ala Gln Ile Gln Ile Val Asn Asp  
 355 360 365  
 Asn Gly Asn Arg Val Phe Asp Gly Gln Val Ser Gln Gly Gln Leu Ile  
 370 375 380  
 Ala Val Pro Gln Gly Phe Ser Val Val Lys Arg Ala Thr Ser Asn Arg  
 385 390 395 400  
 Phe Gln Trp Val Glu Phe Lys Thr Asn Ala Asn Ala Gln Ile Asn Thr  
 405 410 415  
 Leu Ala Gly Arg Thr Ser Val Leu Arg Gly Leu Pro Leu Glu Val Ile  
 420 425 430  
 Thr Asn Gly Phe Gln Ile Ser Pro Glu Glu Ala Arg Arg Val Lys Phe  
 435 440 445  
 Asn Thr Leu Glu Thr Thr Leu Thr His Ser Ser Gly Pro Ala Ser Tyr  
 450 455 460  
 Gly Arg Pro Arg Val Ala Ala Ala  
 465 470

&lt;210&gt; 65

&lt;211&gt; 1368

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1365)

&lt;223&gt; 12S Crb storage protein

&lt;400&gt; 65

atg	ggt	cga	gtc	tca	tct	att	atc	tct	ttc	tct	ttg	aca	ctc	ttg	atc	48
Met	Gly	Arg	Val	Ser	Ser	Ile	Ile	Ser	Phe	Ser	Leu	Thr	Leu	Leu	Ile	
1				5				10					15			
ctc	ttc	aat	ggc	tac	act	gcc	caa	cag	tgg	ccc	aac	gag	tgc	cag	ctc	96
Leu	Phe	Asn	Gly	Tyr	Thr	Ala	Gln	Gln	Trp	Pro	Asn	Glu	Cys	Gln	Leu	
			20					25					30			



## 74

gat	caa	ctc	aat	gcg	ctc	gaa	cca	tcc	caa	atc	atc	aag	agc	gag	ggt	144
Asp	Gln	Leu	Asn	Ala	Leu	Glu	Pro	Ser	Gln	Ile	Ile	Lys	Ser	Glu	Gly	
	35					40						45				
ggt	cgc	atc	gag	gtc	tgg	gac	cac	cat	gca	ccc	cag	ctc	cgt	tgc	tcc	192
Gly	Arg	Ile	Glu	Val	Trp	Asp	His	His	Ala	Pro	Gln	Leu	Arg	Cys	Ser	
	50					55					60					
ggc	ttt	gcc	ttt	gag	cgt	ttc	gtc	att	gag	cct	cag	ggt	ctt	ttc	ttg	240
Gly	Phe	Ala	Phe	Glu	Arg	Phe	Val	Ile	Glu	Pro	Gln	Gly	Leu	Phe	Leu	
	65				70				75						80	
ccc	act	ttc	ttg	aac	gcc	ggc	aaa	ctc	acg	ttt	ggt	ggt	cac	gga	agg	288
Pro	Thr	Phe	Leu	Asn	Ala	Gly	Lys	Leu	Thr	Phe	Val	Val	His	Gly	Arg	
				85					90					95		
ggt	cta	atg	gga	aga	ggt	att	ccg	gga	tgc	gcc	gag	acg	ttc	atg	gag	336
Gly	Leu	Met	Gly	Arg	Val	Ile	Pro	Gly	Cys	Ala	Glu	Thr	Phe	Met	Glu	
			100					105					110			
tca	ccg	gta	ttt	gga	gaa	ggt	caa	ggt	cag	ggt	cag	agt	caa	ggg	ttc	384
Ser	Pro	Val	Phe	Gly	Glu	Gly	Gln	Gly	Gln	Gly	Gln	Ser	Gln	Gly	Phe	
		115					120					125				
cgt	gac	atg	cac	cag	aaa	gta	gag	cac	cta	cgg	tgc	ggt	gac	acc	att	432
Arg	Asp	Met	His	Gln	Lys	Val	Glu	His	Leu	Arg	Cys	Gly	Asp	Thr	Ile	
	130					135					140					
gca	aca	cca	tct	ggt	gta	gct	caa	tgg	ttc	tac	aac	aat	gga	aat	gag	480
Ala	Thr	Pro	Ser	Gly	Val	Ala	Gln	Trp	Phe	Tyr	Asn	Asn	Gly	Asn	Glu	
	145				150				155						160	
cct	ctc	att	ctt	ggt	gca	gcc	gcg	gat	ctc	gcc	agc	aac	cag	aac	cag	528
Pro	Leu	Ile	Leu	Val	Ala	Ala	Ala	Asp	Leu	Ala	Ser	Asn	Gln	Asn	Gln	
				165					170					175		
ctt	gac	cgc	aac	ctt	aga	cca	ttt	ttg	ata	gcc	gga	aac	aac	cca	caa	576
Leu	Asp	Arg	Asn	Leu	Arg	Pro	Phe	Leu	Ile	Ala	Gly	Asn	Asn	Pro	Gln	
			180					185					190			
ggg	cag	gaa	tgg	cta	caa	ggc	cga	aag	caa	cag	aag	caa	aac	aac	atc	624
Gly	Gln	Glu	Trp	Leu	Gln	Gly	Arg	Lys	Gln	Gln	Lys	Gln	Asn	Asn	Ile	
	195						200					205				
ttc	aat	ggc	ttc	gca	cct	gag	atc	ttg	gct	caa	gcc	ttc	aag	atc	aat	672
Phe	Asn	Gly	Phe	Ala	Pro	Glu	Ile	Leu	Ala	Gln	Ala	Phe	Lys	Ile	Asn	
	210					215					220					
gtc	gag	acg	gct	cag	cag	ctc	cag	aac	cag	caa	gat	aac	cgt	ggc	aac	720
Val	Glu	Thr	Ala	Gln	Gln	Leu	Gln	Asn	Gln	Gln	Asp	Asn	Arg	Gly	Asn	
	225				230					235					240	
atc	gtc	aag	gtc	aac	gga	cct	ttc	ggc	gtc	att	agg	cca	ccc	ttg	aga	768
Ile	Val	Lys	Val	Asn	Gly	Pro	Phe	Gly	Val	Ile	Arg	Pro	Pro	Leu	Arg	
				245					250					255		
cgc	ggc	gaa	ggc	ggc	caa	caa	cca	cat	gaa	ata	gct	aat	ggt	tta	gag	816
Arg	Gly	Glu	Gly	Gly	Gln	Gln	Pro	His	Glu	Ile	Ala	Asn	Gly	Leu	Glu	
			260					265					270			
gag	act	ttg	tgc	acc	atg	cga	tgc	act	gaa	aac	ctc	gat	gac	ccg	tcg	864
Glu	Thr	Leu	Cys	Thr	Met	Arg	Cys	Thr	Glu	Asn	Leu	Asp	Asp	Pro	Ser	
		275					280						285			
gat	gct	gac	gtg	tac	aag	cca	tca	ctc	gga	tac	att	agc	aca	ctt	aac	912
Asp	Ala	Asp	Val	Tyr	Lys	Pro	Ser	Leu	Gly	Tyr	Ile	Ser	Thr	Leu	Asn	
	290					295					300					

## 75

```

agc tac aat ctt cct atc ctc aga ctt ctc cgc ctt agc gct ctt cgt      960
Ser Tyr Asn Leu Pro Ile Leu Arg Leu Leu Arg Leu Ser Ala Leu Arg
305                               310                               315                               320

ggc tcc atc cgt aaa aac gct atg gtg cta ccg caa tgg aac gta aac      1008
Gly Ser Ile Arg Lys Asn Ala Met Val Leu Pro Gln Trp Asn Val Asn
                               325                               330                               335

gca aac gcg gca ctc tac gtg aca aac gga aag gct cat ata caa atg      1056
Ala Asn Ala Ala Leu Tyr Val Thr Asn Gly Lys Ala His Ile Gln Met
                               340                               345                               350

gtg aac gac aac gga gaa aga gtg ttc gac caa gag atc tcc agc gga      1104
Val Asn Asp Asn Gly Glu Arg Val Phe Asp Gln Glu Ile Ser Ser Gly
                               355                               360                               365

cag tta cta gtc gtg cca caa ggc ttt tgc gtc atg aaa cat cgc ata      1152
Gln Leu Leu Val Val Pro Gln Gly Phe Ser Val Met Lys His Arg Ile
                               370                               375                               380

ggc gaa cag ttc gag tgg atc gaa ttc aag aca aac gaa aac gca cag      1200
Gly Glu Gln Phe Glu Trp Ile Glu Phe Lys Thr Asn Glu Asn Ala Gln
385                               390                               395                               400

gtc aac aca ctc gcg ggc cgt acc tca gtc atg aga ggt ttg ccg ctt      1248
Val Asn Thr Leu Ala Gly Arg Thr Ser Val Met Arg Gly Leu Pro Leu
                               405                               410                               415

gag gtt ata acc aat ggg tac cag atc tct ccc gaa gaa gct aaa cga      1296
Glu Val Ile Thr Asn Gly Tyr Gln Ile Ser Pro Glu Glu Ala Lys Arg
                               420                               425                               430

gta aag ttt agc acg att gag acc aca ctg acc cat agc agt cca atg      1344
Val Lys Phe Ser Thr Ile Glu Thr Thr Leu Thr His Ser Ser Pro Met
                               435                               440                               445

agc tac gga agg cct agg gct tga                                     1368
Ser Tyr Gly Arg Pro Arg Ala
                               450                               455

<210> 66
<211> 455
<212> PRT
<213> Arabidopsis thaliana
<400> 66
Met Gly Arg Val Ser Ser Ile Ile Ser Phe Ser Leu Thr Leu Leu Ile
 1                               5                               10                               15
Leu Phe Asn Gly Tyr Thr Ala Gln Gln Trp Pro Asn Glu Cys Gln Leu
                               20                               25                               30
Asp Gln Leu Asn Ala Leu Glu Pro Ser Gln Ile Ile Lys Ser Glu Gly
                               35                               40                               45
Gly Arg Ile Glu Val Trp Asp His His Ala Pro Gln Leu Arg Cys Ser
                               50                               55                               60
Gly Phe Ala Phe Glu Arg Phe Val Ile Glu Pro Gln Gly Leu Phe Leu
65                               70                               75                               80
Pro Thr Phe Leu Asn Ala Gly Lys Leu Thr Phe Val Val His Gly Arg
                               85                               90                               95
Gly Leu Met Gly Arg Val Ile Pro Gly Cys Ala Glu Thr Phe Met Glu
                               100                              105                              110
Ser Pro Val Phe Gly Glu Gly Gln Gly Gln Gly Gln Ser Gln Gly Phe
                               115                              120                              125

```

## 76

Arg Asp Met His Gln Lys Val Glu His Leu Arg Cys Gly Asp Thr Ile  
 130 135 140  
 Ala Thr Pro Ser Gly Val Ala Gln Trp Phe Tyr Asn Asn Gly Asn Glu  
 145 150 155 160  
 Pro Leu Ile Leu Val Ala Ala Ala Asp Leu Ala Ser Asn Gln Asn Gln  
 165 170 175  
 Leu Asp Arg Asn Leu Arg Pro Phe Leu Ile Ala Gly Asn Asn Pro Gln  
 180 185 190  
 Gly Gln Glu Trp Leu Gln Gly Arg Lys Gln Gln Lys Gln Asn Asn Ile  
 195 200 205  
 Phe Asn Gly Phe Ala Pro Glu Ile Leu Ala Gln Ala Phe Lys Ile Asn  
 210 215 220  
 Val Glu Thr Ala Gln Gln Leu Gln Asn Gln Gln Asp Asn Arg Gly Asn  
 225 230 235 240  
 Ile Val Lys Val Asn Gly Pro Phe Gly Val Ile Arg Pro Pro Leu Arg  
 245 250 255  
 Arg Gly Glu Gly Gly Gln Gln Pro His Glu Ile Ala Asn Gly Leu Glu  
 260 265 270  
 Glu Thr Leu Cys Thr Met Arg Cys Thr Glu Asn Leu Asp Asp Pro Ser  
 275 280 285  
 Asp Ala Asp Val Tyr Lys Pro Ser Leu Gly Tyr Ile Ser Thr Leu Asn  
 290 295 300  
 Ser Tyr Asn Leu Pro Ile Leu Arg Leu Leu Arg Leu Ser Ala Leu Arg  
 305 310 315 320  
 Gly Ser Ile Arg Lys Asn Ala Met Val Leu Pro Gln Trp Asn Val Asn  
 325 330 335  
 Ala Asn Ala Ala Leu Tyr Val Thr Asn Gly Lys Ala His Ile Gln Met  
 340 345 350  
 Val Asn Asp Asn Gly Glu Arg Val Phe Asp Gln Glu Ile Ser Ser Gly  
 355 360 365  
 Gln Leu Leu Val Val Pro Gln Gly Phe Ser Val Met Lys His Arg Ile  
 370 375 380  
 Gly Glu Gln Phe Glu Trp Ile Glu Phe Lys Thr Asn Glu Asn Ala Gln  
 385 390 395 400  
 Val Asn Thr Leu Ala Gly Arg Thr Ser Val Met Arg Gly Leu Pro Leu  
 405 410 415  
 Glu Val Ile Thr Asn Gly Tyr Gln Ile Ser Pro Glu Glu Ala Lys Arg  
 420 425 430  
 Val Lys Phe Ser Thr Ile Glu Thr Thr Leu Thr His Ser Ser Pro Met  
 435 440 445  
 Ser Tyr Gly Arg Pro Arg Ala  
 450 455

&lt;210&gt; 67

&lt;211&gt; 1356

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1353)

&lt;223&gt; putative 12S storage protein

&lt;400&gt; 67

atg cat aag ctt ttg ttt tct ctt ctc tcc gtc gtc tca ctc tca ttt	48
Met His Lys Leu Leu Phe Ser Leu Leu Ser Val Val Ser Leu Ser Phe	
1 5 10 15	
ctc ctc ttc ttc cat ggc gcc gag gca cgc cag cga gag gcg ccg ttt	96
Leu Leu Phe Phe His Gly Ala Glu Ala Arg Gln Arg Glu Ala Pro Phe	
20 25 30	
cca aac gcc tgc cat ttc agc caa atc aac agc ctc gcg ccc gct cag	144
Pro Asn Ala Cys His Phe Ser Gln Ile Asn Ser Leu Ala Pro Ala Gln	
35 40 45	
gcg acg aag ttc gaa gcc ggt cag atg gaa gta tgg gac cac atg agc	192
Ala Thr Lys Phe Glu Ala Gly Gln Met Glu Val Trp Asp His Met Ser	
50 55 60	
cct gag ctc cga tgc gcc ggt gta acg gtg gct cgc atc acc ctt cag	240
Pro Glu Leu Arg Cys Ala Gly Val Thr Val Ala Arg Ile Thr Leu Gln	
65 70 75 80	
ccc aat tcc att ttc ttg ccc gct ttc ttt agc cca cct gcc ctt gct	288
Pro Asn Ser Ile Phe Leu Pro Ala Phe Phe Ser Pro Pro Ala Leu Ala	
85 90 95	
tac gtt gtc caa gga gaa gga gtt atg ggg acg att gct tct ggt tgt	336
Tyr Val Val Gln Gly Glu Gly Val Met Gly Thr Ile Ala Ser Gly Cys	
100 105 110	
cct gag act ttt gca gaa gtt gaa gga tca tca gga aga gga gga gga	384
Pro Glu Thr Phe Ala Glu Val Glu Gly Ser Ser Gly Arg Gly Gly Gly	
115 120 125	
gga gac ccg ggt cga cgt ttt gag gac atg cac cag aag ttg gag aat	432
Gly Asp Pro Gly Arg Arg Phe Glu Asp Met His Gln Lys Leu Glu Asn	
130 135 140	
ttc cgg cga ggg gat gtg ttt gct tcg ctt gcc gga gtt tca cag tgg	480
Phe Arg Arg Gly Asp Val Phe Ala Ser Leu Ala Gly Val Ser Gln Trp	
145 150 155 160	
tgg tac aac cgc ggt gat tcc gat gcc gtc att gtc att gtt ctt gat	528
Trp Tyr Asn Arg Gly Asp Ser Asp Ala Val Ile Val Ile Val Leu Asp	
165 170 175	
gtc acc aac aga gaa aac cag ctt gac caa gtc cct agg atg ttc caa	576
Val Thr Asn Arg Glu Asn Gln Leu Asp Gln Val Pro Arg Met Phe Gln	
180 185 190	
cta gcc ggg agc aga acg caa gaa gaa gaa caa cca tta acg tgg cca	624
Leu Ala Gly Ser Arg Thr Gln Glu Glu Glu Gln Pro Leu Thr Trp Pro	
195 200 205	
tca ggc aac aac gct ttc agc ggt ttc gac cca aac ata atc gcg gaa	672
Ser Gly Asn Asn Ala Phe Ser Gly Phe Asp Pro Asn Ile Ile Ala Glu	
210 215 220	
gca ttc aaa atc aac atc gag aca gct aag caa cta caa aac cag aag	720
Ala Phe Lys Ile Asn Ile Glu Thr Ala Lys Gln Leu Gln Asn Gln Lys	
225 230 235 240	
gac aac aga gga aac ata atc cga gca aat ggt cct ctc cat ttc gtc	768
Asp Asn Arg Gly Asn Ile Ile Arg Ala Asn Gly Pro Leu His Phe Val	
245 250 255	

## 78

atc cca ccg cct cgt gaa tgg cag caa gat ggc att gct aat ggc atc 816  
 Ile Pro Pro Pro Arg Glu Trp Gln Gln Asp Gly Ile Ala Asn Gly Ile  
 260 265 270

gaa gag act tat tgc acg gct aag att cat gag aat atc gat gat cca 864  
 Glu Glu Thr Tyr Cys Thr Ala Lys Ile His Glu Asn Ile Asp Asp Pro  
 275 280 285

gaa cgg tct gac cat ttt agc aca cga gcc gga aga atc agc act ctt 912  
 Glu Arg Ser Asp His Phe Ser Thr Arg Ala Gly Arg Ile Ser Thr Leu  
 290 295 300

aac agc ctt aat ctc cct gtt cta cgt cta gtc aga ctt aac gcc ctt 960  
 Asn Ser Leu Asn Leu Pro Val Leu Arg Leu Val Arg Leu Asn Ala Leu  
 305 310 315 320

aga ggt tat ctc tac agc gga gga atg gtg ttg cca caa tgg acg gca 1008  
 Arg Gly Tyr Leu Tyr Ser Gly Gly Met Val Leu Pro Gln Trp Thr Ala  
 325 330 335

aac gcg cac acg gtg cta tac gtc aca gga ggt caa gcc aag ata caa 1056  
 Asn Ala His Thr Val Leu Tyr Val Thr Gly Gly Gln Ala Lys Ile Gln  
 340 345 350

gtg gtg gac gac aat ggt cag tcg gtg ttc aat gag caa gtg gga caa 1104  
 Val Val Asp Asp Asn Gly Gln Ser Val Phe Asn Glu Gln Val Gly Gln  
 355 360 365

ggc caa atc att gtg att cca caa ggc ttt gca gtt tca aaa acg gct 1152  
 Gly Gln Ile Ile Val Ile Pro Gln Gly Phe Ala Val Ser Lys Thr Ala  
 370 375 380

ggt gaa acg ggt ttc gag tgg ata tca ttc aag aca aac gat aac gct 1200  
 Gly Glu Thr Gly Phe Glu Trp Ile Ser Phe Lys Thr Asn Asp Asn Ala  
 385 390 395 400

tac att aac aca ctg agc ggc caa aca tcg tac ttg aga gca gtt cca 1248  
 Tyr Ile Asn Thr Leu Ser Gly Gln Thr Ser Tyr Leu Arg Ala Val Pro  
 405 410 415

gtg gat gtg atc aaa gcg tca tat gga gtg aac gag gaa gaa gcc aag 1296  
 Val Asp Val Ile Lys Ala Ser Tyr Gly Val Asn Glu Glu Glu Ala Lys  
 420 425 430

agg atc aag ttt agt cag caa gag acc atg ttg tct atg aca cca agc 1344  
 Arg Ile Lys Phe Ser Gln Gln Glu Thr Met Leu Ser Met Thr Pro Ser  
 435 440 445

tct tct tct taa 1356  
 Ser Ser Ser  
 450

<210> 68  
 <211> 451  
 <212> PRT  
 <213> Arabidopsis thaliana

<400> 68  
 Met His Lys Leu Leu Phe Ser Leu Leu Ser Val Val Ser Leu Ser Phe  
 1 5 10 15  
 Leu Leu Phe Phe His Gly Ala Glu Ala Arg Gln Arg Glu Ala Pro Phe  
 20 25 30  
 Pro Asn Ala Cys His Phe Ser Gln Ile Asn Ser Leu Ala Pro Ala Gln  
 35 40 45

Ala	Thr	Lys	Phe	Glu	Ala	Gly	Gln	Met	Glu	Val	Trp	Asp	His	Met	Ser	50	55	60
Pro	Glu	Leu	Arg	Cys	Ala	Gly	Val	Thr	Val	Ala	Arg	Ile	Thr	Leu	Gln	65	70	75
Pro	Asn	Ser	Ile	Phe	Leu	Pro	Ala	Phe	Phe	Ser	Pro	Pro	Ala	Leu	Ala	85	90	95
Tyr	Val	Val	Gln	Gly	Glu	Gly	Val	Met	Gly	Thr	Ile	Ala	Ser	Gly	Cys	100	105	110
Pro	Glu	Thr	Phe	Ala	Glu	Val	Glu	Gly	Ser	Ser	Gly	Arg	Gly	Gly	Gly	115	120	125
Gly	Asp	Pro	Gly	Arg	Arg	Phe	Glu	Asp	Met	His	Gln	Lys	Leu	Glu	Asn	130	135	140
Phe	Arg	Arg	Gly	Asp	Val	Phe	Ala	Ser	Leu	Ala	Gly	Val	Ser	Gln	Trp	145	150	155
Trp	Tyr	Asn	Arg	Gly	Asp	Ser	Asp	Ala	Val	Ile	Val	Ile	Val	Leu	Asp	165	170	175
Val	Thr	Asn	Arg	Glu	Asn	Gln	Leu	Asp	Gln	Val	Pro	Arg	Met	Phe	Gln	180	185	190
Leu	Ala	Gly	Ser	Arg	Thr	Gln	Glu	Glu	Gln	Pro	Leu	Thr	Trp	Pro		195	200	205
Ser	Gly	Asn	Asn	Ala	Phe	Ser	Gly	Phe	Asp	Pro	Asn	Ile	Ile	Ala	Glu	210	215	220
Ala	Phe	Lys	Ile	Asn	Ile	Glu	Thr	Ala	Lys	Gln	Leu	Gln	Asn	Gln	Lys	225	230	235
Asp	Asn	Arg	Gly	Asn	Ile	Ile	Arg	Ala	Asn	Gly	Pro	Leu	His	Phe	Val	245	250	255
Ile	Pro	Pro	Pro	Arg	Glu	Trp	Gln	Gln	Asp	Gly	Ile	Ala	Asn	Gly	Ile	260	265	270
Glu	Glu	Thr	Tyr	Cys	Thr	Ala	Lys	Ile	His	Glu	Asn	Ile	Asp	Asp	Pro	275	280	285
Glu	Arg	Ser	Asp	His	Phe	Ser	Thr	Arg	Ala	Gly	Arg	Ile	Ser	Thr	Leu	290	295	300
Asn	Ser	Leu	Asn	Leu	Pro	Val	Leu	Arg	Leu	Val	Arg	Leu	Asn	Ala	Leu	305	310	315
Arg	Gly	Tyr	Leu	Tyr	Ser	Gly	Gly	Met	Val	Leu	Pro	Gln	Trp	Thr	Ala	325	330	335
Asn	Ala	His	Thr	Val	Leu	Tyr	Val	Thr	Gly	Gly	Gln	Ala	Lys	Ile	Gln	340	345	350
Val	Val	Asp	Asp	Asn	Gly	Gln	Ser	Val	Phe	Asn	Glu	Gln	Val	Gly	Gln	355	360	365
Gly	Gln	Ile	Ile	Val	Ile	Pro	Gln	Gly	Phe	Ala	Val	Ser	Lys	Thr	Ala	370	375	380
Gly	Glu	Thr	Gly	Phe	Glu	Trp	Ile	Ser	Phe	Lys	Thr	Asn	Asp	Asn	Ala	385	390	395
Tyr	Ile	Asn	Thr	Leu	Ser	Gly	Gln	Thr	Ser	Tyr	Leu	Arg	Ala	Val	Pro	405	410	415
Val	Asp	Val	Ile	Lys	Ala	Ser	Tyr	Gly	Val	Asn	Glu	Glu	Glu	Ala	Lys	420	425	430

80

Arg Ile Lys Phe Ser Gln Gln Glu Thr Met Leu Ser Met Thr Pro Ser  
 435 440 445  
 Ser Ser Ser  
 450

&lt;210&gt; 69

&lt;211&gt; 1356

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1353)

&lt;223&gt; 12S storage protein At1g03890

&lt;400&gt; 69

atg cat aag ctt ttg ttt tct ctt ctc tcc gtc gtc tca ctc tca ttt	48
Met His Lys Leu Leu Phe Ser Leu Leu Ser Val Val Ser Leu Ser Phe	
1 5 10 15	
ctc ctc ttc ttc cat ggc gcc gag gca cgc cag cga gag gcg ccg ttt	96
Leu Leu Phe Phe His Gly Ala Glu Ala Arg Gln Arg Glu Ala Pro Phe	
20 25 30	
cca aac gcc tgc cat ttc agc caa atc aac agc ctc gcg ccc gct cag	144
Pro Asn Ala Cys His Phe Ser Gln Ile Asn Ser Leu Ala Pro Ala Gln	
35 40 45	
gcg acg aag ttc gaa gcc ggt cag atg gaa gta tgg gac cac atg agc	192
Ala Thr Lys Phe Glu Ala Gly Gln Met Glu Val Trp Asp His Met Ser	
50 55 60	
cct gag ctc cga tgc gcc ggt gta acg gtg gct cgc atc acc ctt cag	240
Pro Glu Leu Arg Cys Ala Gly Val Thr Val Ala Arg Ile Thr Leu Gln	
65 70 75 80	
ccc aat tcc att ttc ttg ccc gct ttc ttt agc cca cct gcc ctt gct	288
Pro Asn Ser Ile Phe Leu Pro Ala Phe Phe Ser Pro Pro Ala Leu Ala	
85 90 95	
tac gtt gtc caa gga gaa gga gtt atg ggg acg att gct tct ggt tgt	336
Tyr Val Val Gln Gly Glu Gly Val Met Gly Thr Ile Ala Ser Gly Cys	
100 105 110	
cct gag act ttt gca gaa gtt gaa gga tca tca gga aga gga gga gga	384
Pro Glu Thr Phe Ala Glu Val Glu Gly Ser Ser Gly Arg Gly Gly Gly	
115 120 125	
gga gac ccg ggt cga cgt ttt gag gac atg cac cag aag ttg gag aat	432
Gly Asp Pro Gly Arg Arg Phe Glu Asp Met His Gln Lys Leu Glu Asn	
130 135 140	
ttc cgg cga ggg gat gtg ttt gct tcg ctt gcc gga gtt tca cag tgg	480
Phe Arg Arg Gly Asp Val Phe Ala Ser Leu Ala Gly Val Ser Gln Trp	
145 150 155 160	
tgg tac aac cgc ggt gat tcc gat gcc gtc att gtc att gtt ctt gat	528
Trp Tyr Asn Arg Gly Asp Ser Asp Ala Val Ile Val Ile Val Leu Asp	
165 170 175	
gtc acc aac aga gaa aac cag ctt gac caa gtc cct agg atg ttc caa	576
Val Thr Asn Arg Glu Asn Gln Leu Asp Gln Val Pro Arg Met Phe Gln	
180 185 190	

cta gcc ggg agc aga acg caa gaa gaa gaa caa cca tta acg tgg cca	624
Leu Ala Gly Ser Arg Thr Gln Glu Glu Glu Gln Pro Leu Thr Trp Pro	
195 200 205	
tca ggc aac aac gct ttc agc ggt ttc gac cca aac ata atc gcg gaa	672
Ser Gly Asn Asn Ala Phe Ser Gly Phe Asp Pro Asn Ile Ile Ala Glu	
210 215 220	
gca ttc aaa atc aac atc gag aca gct aag caa cta caa aac cag aag	720
Ala Phe Lys Ile Asn Ile Glu Thr Ala Lys Gln Leu Gln Asn Gln Lys	
225 230 235 240	
gac aac aga gga aac ata atc cga gca aat ggt cct ctc cat ttc gtc	768
Asp Asn Arg Gly Asn Ile Ile Arg Ala Asn Gly Pro Leu His Phe Val	
245 250 255	
atc cca ccg cct cgt gaa tgg cag caa gat ggc att gct aat ggc atc	816
Ile Pro Pro Pro Arg Glu Trp Gln Gln Asp Gly Ile Ala Asn Gly Ile	
260 265 270	
gaa gag act tat tgc acg gct aag att cat gag aat atc gat gat cca	864
Glu Glu Thr Tyr Cys Thr Ala Lys Ile His Glu Asn Ile Asp Asp Pro	
275 280 285	
gaa cgg tct gac cat ttt agc aca cga gcc gga aga atc agc act ctt	912
Glu Arg Ser Asp His Phe Ser Thr Arg Ala Gly Arg Ile Ser Thr Leu	
290 295 300	
aac agc ctt aat ctc cct gtt cta cgt cta gtc aga ctt aac gcc ctt	960
Asn Ser Leu Asn Leu Pro Val Leu Arg Leu Val Arg Leu Asn Ala Leu	
305 310 315 320	
aga ggt tat ctc tac agc gga gga atg gtg ttg cca caa tgg acg gca	1008
Arg Gly Tyr Leu Tyr Ser Gly Gly Met Val Leu Pro Gln Trp Thr Ala	
325 330 335	
aac gcg cac acg gtg cta tac gtc aca gga ggt caa gcc aag ata caa	1056
Asn Ala His Thr Val Leu Tyr Val Thr Gly Gly Gln Ala Lys Ile Gln	
340 345 350	
gtg gtg gac gac aat ggt cag tcg gtg ttc aat gag caa gtg gga caa	1104
Val Val Asp Asp Asn Gly Gln Ser Val Phe Asn Glu Gln Val Gly Gln	
355 360 365	
ggc caa atc att gtg att cca caa ggc ttt gca gtt tca aaa acg gct	1152
Gly Gln Ile Ile Val Ile Pro Gln Gly Phe Ala Val Ser Lys Thr Ala	
370 375 380	
ggt gaa acg ggt ttc gag tgg ata tca ttc aag aca aac gat aac gct	1200
Gly Glu Thr Gly Phe Glu Trp Ile Ser Phe Lys Thr Asn Asp Asn Ala	
385 390 395 400	
tac att aac aca ctg agc ggc caa aca tcg tac ttg aga gca gtt cca	1248
Tyr Ile Asn Thr Leu Ser Gly Gln Thr Ser Tyr Leu Arg Ala Val Pro	
405 410 415	
gtg gat gtg atc aaa gcg tca tat gga gtg aac gag gaa gaa gcc aag	1296
Val Asp Val Ile Lys Ala Ser Tyr Gly Val Asn Glu Glu Glu Ala Lys	
420 425 430	
agg atc aag ttt agt cag caa gag acc atg ttg tct atg aca cca agc	1344
Arg Ile Lys Phe Ser Gln Gln Glu Thr Met Leu Ser Met Thr Pro Ser	
435 440 445	
tct tct tct taa	1356
Ser Ser Ser	
450	



&lt;210&gt; 70

&lt;211&gt; 451

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 70

```

Met His Lys Leu Leu Phe Ser Leu Leu Ser Val Val Ser Leu Ser Phe
 1           5           10           15
Leu Leu Phe Phe His Gly Ala Glu Ala Arg Gln Arg Glu Ala Pro Phe
          20           25           30
Pro Asn Ala Cys His Phe Ser Gln Ile Asn Ser Leu Ala Pro Ala Gln
          35           40           45
Ala Thr Lys Phe Glu Ala Gly Gln Met Glu Val Trp Asp His Met Ser
          50           55           60
Pro Glu Leu Arg Cys Ala Gly Val Thr Val Ala Arg Ile Thr Leu Gln
          65           70           75           80
Pro Asn Ser Ile Phe Leu Pro Ala Phe Phe Ser Pro Pro Ala Leu Ala
          85           90           95
Tyr Val Val Gln Gly Glu Gly Val Met Gly Thr Ile Ala Ser Gly Cys
          100          105          110
Pro Glu Thr Phe Ala Glu Val Glu Gly Ser Ser Gly Arg Gly Gly Gly
          115          120          125
Gly Asp Pro Gly Arg Arg Phe Glu Asp Met His Gln Lys Leu Glu Asn
          130          135          140
Phe Arg Arg Gly Asp Val Phe Ala Ser Leu Ala Gly Val Ser Gln Trp
          145          150          155          160
Trp Tyr Asn Arg Gly Asp Ser Asp Ala Val Ile Val Ile Val Leu Asp
          165          170          175
Val Thr Asn Arg Glu Asn Gln Leu Asp Gln Val Pro Arg Met Phe Gln
          180          185          190
Leu Ala Gly Ser Arg Thr Gln Glu Glu Gln Pro Leu Thr Trp Pro
          195          200          205
Ser Gly Asn Asn Ala Phe Ser Gly Phe Asp Pro Asn Ile Ile Ala Glu
          210          215          220
Ala Phe Lys Ile Asn Ile Glu Thr Ala Lys Gln Leu Gln Asn Gln Lys
          225          230          235          240
Asp Asn Arg Gly Asn Ile Ile Arg Ala Asn Gly Pro Leu His Phe Val
          245          250          255
Ile Pro Pro Pro Arg Glu Trp Gln Gln Asp Gly Ile Ala Asn Gly Ile
          260          265          270
Glu Glu Thr Tyr Cys Thr Ala Lys Ile His Glu Asn Ile Asp Asp Pro
          275          280          285
Glu Arg Ser Asp His Phe Ser Thr Arg Ala Gly Arg Ile Ser Thr Leu
          290          295          300
Asn Ser Leu Asn Leu Pro Val Leu Arg Leu Val Arg Leu Asn Ala Leu
          305          310          315          320
Arg Gly Tyr Leu Tyr Ser Gly Gly Met Val Leu Pro Gln Trp Thr Ala
          325          330          335
Asn Ala His Thr Val Leu Tyr Val Thr Gly Gly Gln Ala Lys Ile Gln
          340          345          350

```

## 83

Val Val Asp Asp Asn Gly Gln Ser Val Phe Asn Glu Gln Val Gly Gln  
 355 360 365  
 Gly Gln Ile Ile Val Ile Pro Gln Gly Phe Ala Val Ser Lys Thr Ala  
 370 375 380  
 Gly Glu Thr Gly Phe Glu Trp Ile Ser Phe Lys Thr Asn Asp Asn Ala  
 385 390 395 400  
 Tyr Ile Asn Thr Leu Ser Gly Gln Thr Ser Tyr Leu Arg Ala Val Pro  
 405 410 415  
 Val Asp Val Ile Lys Ala Ser Tyr Gly Val Asn Glu Glu Glu Ala Lys  
 420 425 430  
 Arg Ile Lys Phe Ser Gln Gln Glu Thr Met Leu Ser Met Thr Pro Ser  
 435 440 445  
 Ser Ser Ser  
 450

&lt;210&gt; 71

&lt;211&gt; 867

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(864)

&lt;223&gt; prohibitin 1

&lt;400&gt; 71

atg aac aac gtc aaa gtt cca aag ata cca ggt ggt ggt gcc att tcg 48  
 Met Asn Asn Val Lys Val Pro Lys Ile Pro Gly Gly Gly Ala Ile Ser  
 1 5 10 15  
 acg ttg ctt aag gtt ggg att att ggt ggg ctt ggc ctc tat ggt gct 96  
 Thr Leu Leu Lys Val Gly Ile Ile Gly Gly Leu Gly Leu Tyr Gly Ala  
 20 25 30  
 acg cac agt ctc tac aat gtt gaa gga gga cat cga gcc atc atg ttc 144  
 Thr His Ser Leu Tyr Asn Val Glu Gly Gly His Arg Ala Ile Met Phe  
 35 40 45  
 aat cgt tta gtc ggt att aaa gat aag gtt tac cct gag ggt aca cac 192  
 Asn Arg Leu Val Gly Ile Lys Asp Lys Val Tyr Pro Glu Gly Thr His  
 50 55 60  
 ctt atg att cct tgg ttt gaa agg ccg gtc atc tat gac gtt cgt gct 240  
 Leu Met Ile Pro Trp Phe Glu Arg Pro Val Ile Tyr Asp Val Arg Ala  
 65 70 75 80  
 cga cct tac ctt gtt gag agt aca tcc gga agc cgt gat ctt cag atg 288  
 Arg Pro Tyr Leu Val Glu Ser Thr Ser Gly Ser Arg Asp Leu Gln Met  
 85 90 95  
 gtg aaa att ggg ctt agg gtt ctc aca cgt ccc atg gca gac cag tta 336  
 Val Lys Ile Gly Leu Arg Val Leu Thr Arg Pro Met Ala Asp Gln Leu  
 100 105 110  
 cct gaa atc tac aga agc ctt ggt gag aac tac agc gag aga gtt cta 384  
 Pro Glu Ile Tyr Arg Ser Leu Gly Glu Asn Tyr Ser Glu Arg Val Leu  
 115 120 125  
 cct tct ata atc aac gag act ttg aaa gct gtg gtt gct cag tac aat 432  
 Pro Ser Ile Ile Asn Glu Thr Leu Lys Ala Val Val Ala Gln Tyr Asn  
 130 135 140

84

gca agc cag ctt att act cag aga gag gcg gtc agt agg gag atc agg 480  
 Ala Ser Gln Leu Ile Thr Gln Arg Glu Ala Val Ser Arg Glu Ile Arg  
 145 150 155 160  
 aag att ctg act gaa cga gca gca aac ttc aat gtt gcg ctt gac gat 528  
 Lys Ile Leu Thr Glu Arg Ala Ala Asn Phe Asn Val Ala Leu Asp Asp  
 165 170 175  
 gtg tcc atc aca aac ctg aca ttc ggg aag gag ttc aca gct gcc att 576  
 Val Ser Ile Thr Asn Leu Thr Phe Gly Lys Glu Phe Thr Ala Ala Ile  
 180 185 190  
 gaa gca aag cag gtg gcg gct caa gag gct gag cgg gct aag ttc att 624  
 Glu Ala Lys Gln Val Ala Ala Gln Glu Ala Glu Arg Ala Lys Phe Ile  
 195 200 205  
 gtt gag aag gcc gaa caa gac aag aga agt gct gtt atc cgc gcc cag 672  
 Val Glu Lys Ala Glu Gln Asp Lys Arg Ser Ala Val Ile Arg Ala Gln  
 210 215 220  
 gga gaa gcc aag agt gct cag ctc att ggt caa gca att gca aac aac 720  
 Gly Glu Ala Lys Ser Ala Gln Leu Ile Gly Gln Ala Ile Ala Asn Asn  
 225 230 235 240  
 caa gcg ttt atc acg ctc agg aag atc gag gct gca aga gag att gca 768  
 Gln Ala Phe Ile Thr Leu Arg Lys Ile Glu Ala Ala Arg Glu Ile Ala  
 245 250 255  
 cag acc ata gca aac tcg gcg aac aag gtt tac ttg agc tca gac gat 816  
 Gln Thr Ile Ala Asn Ser Ala Asn Lys Val Tyr Leu Ser Ser Asp Asp  
 260 265 270  
 ctt ttg ctt aac cta caa ggg atg aat ttg gat gtt gat gca aag aac 864  
 Leu Leu Leu Asn Leu Gln Gly Met Asn Leu Asp Val Asp Ala Lys Asn  
 275 280 285  
 tag 867  
 <210> 72  
 <211> 288  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 72  
 Met Asn Asn Val Lys Val Pro Lys Ile Pro Gly Gly Gly Ala Ile Ser  
 1 5 10 15  
 Thr Leu Leu Lys Val Gly Ile Ile Gly Gly Leu Gly Leu Tyr Gly Ala  
 20 25 30  
 Thr His Ser Leu Tyr Asn Val Glu Gly Gly His Arg Ala Ile Met Phe  
 35 40 45  
 Asn Arg Leu Val Gly Ile Lys Asp Lys Val Tyr Pro Glu Gly Thr His  
 50 55 60  
 Leu Met Ile Pro Trp Phe Glu Arg Pro Val Ile Tyr Asp Val Arg Ala  
 65 70 75 80  
 Arg Pro Tyr Leu Val Glu Ser Thr Ser Gly Ser Arg Asp Leu Gln Met  
 85 90 95  
 Val Lys Ile Gly Leu Arg Val Leu Thr Arg Pro Met Ala Asp Gln Leu  
 100 105 110  
 Pro Glu Ile Tyr Arg Ser Leu Gly Glu Asn Tyr Ser Glu Arg Val Leu  
 115 120 125

## 85

Pro Ser Ile Ile Asn Glu Thr Leu Lys Ala Val Val Ala Gln Tyr Asn  
 130 135 140  
 Ala Ser Gln Leu Ile Thr Gln Arg Glu Ala Val Ser Arg Glu Ile Arg  
 145 150 155 160  
 Lys Ile Leu Thr Glu Arg Ala Ala Asn Phe Asn Val Ala Leu Asp Asp  
 165 170 175  
 Val Ser Ile Thr Asn Leu Thr Phe Gly Lys Glu Phe Thr Ala Ala Ile  
 180 185 190  
 Glu Ala Lys Gln Val Ala Ala Gln Glu Ala Glu Arg Ala Lys Phe Ile  
 195 200 205  
 Val Glu Lys Ala Glu Gln Asp Lys Arg Ser Ala Val Ile Arg Ala Gln  
 210 215 220  
 Gly Glu Ala Lys Ser Ala Gln Leu Ile Gly Gln Ala Ile Ala Asn Asn  
 225 230 235 240  
 Gln Ala Phe Ile Thr Leu Arg Lys Ile Glu Ala Ala Arg Glu Ile Ala  
 245 250 255  
 Gln Thr Ile Ala Asn Ser Ala Asn Lys Val Tyr Leu Ser Ser Asp Asp  
 260 265 270  
 Leu Leu Leu Asn Leu Gln Gly Met Asn Leu Asp Val Asp Ala Lys Asn  
 275 280 285

&lt;210&gt; 73

&lt;211&gt; 40

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotid primer

&lt;400&gt; 73

ataagaatgc ggccgcgtgt tccatttggc cggaacaac

40

&lt;210&gt; 74

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotid primer

&lt;400&gt; 74

cccggatcct tctgtaacat ttgacaaaac atg

33

&lt;210&gt; 75

&lt;211&gt; 40

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotid primer

&lt;400&gt; 75

ataagaatgc ggccgcgtgt tccatttggc cggaacaac

40

&lt;210&gt; 76

&lt;211&gt; 44

<212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotid primer  
 <400> 76  
 ataagaatgc ggccgcggat ccaccctgga gaacgccacg agtg 44  
 <210> 77  
 <211> 45  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotid primer  
 <400> 77  
 ataagaatgc ggccgcggat ccctcagggt cttttcttgc ccact 45  
 <210> 78  
 <211> 30  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotid primer  
 <400> 78  
 ccgctcgagt ttacggatgg agccacgaag 30  
 <210> 79  
 <211> 30  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotid primer  
 <400> 79  
 ccgctcgagg taagctcaac aaatctttag 30  
 <210> 80  
 <211> 31  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotid primer  
 <400> 80  
 acgcgtcgac gcgttctgcg tgcaagatat t 31  
 <210> 81  
 <211> 46  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
         oligonucleotid primer

<400> 81  
 ataagaatgc ggccgcggat ccatggctaa caagctcttc ctgcgc 46

<210> 82  
 <211> 45  
 <212> DNA  
 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz:  
 oligonucleotid primer

<400> 82  
 ataagaatgc ggccgcggat ccctagtagt aaggaggga gaaag 45

<210> 83  
 <211> 4954  
 <212> DNA  
 <213> Künstliche Sequenz

<220>  
 <223> Beschreibung der künstlichen Sequenz: DNA  
 construct coding for dsRNA for suppression of  
 multiple storage proteins

<400> 83  
 agcttggtac cgagctcggg tccactagta acggccgcca gtgtgctgga attcgccctt 60  
 gcggccgcgt gttccatttg gccggaaca accagcaggg aggctttggc ggttcacagc 120  
 aacaacaaga acagaaaaac ttgtggagcg gggttcgacgc acaggtcata gctcaagcat 180  
 tgaaaattga cgttcagttg gctcagcagc ttcagaacca acaagacagc agaggaaaca 240  
 tcgttcgtgt taagggaacct ttccagggtcg tgaggccacc tctaagacag ccctacgaga 300  
 gcgaggagtg gagacaccca cgtagccacc agggcaacgg ccttgaggag cctatctgca 360  
 gcatgaggtc ccacgagaac attgacgacc ctgctcgtgc tgacgtgtac aagcccagcc 420  
 taggtcgcgt gaccagcgtc aacagctata ccttgcccat cttggagtat gtcagggtca 480  
 gtgccactcg tggcgttctc cagggtggat ccttctgtaa catttgacaa aacatgtgaa 540  
 cagtcacatc gtcatataga acttccaatt ttaatatgtt ttgctaaaga aaaaaaaaaag 600  
 gaataaatat ctatcaaatt cattttttaa acatttgtat acgttcttaa ataatttagg 660  
 atatgactaa tttttctttt tggtaaaaat gttaatatct atattttaatt tattaagaaa 720  
 aatgtactta caccctggag aacgccacga gtggcactga gcctgacata ctccaagatg 780  
 ggcaagggtat agctgttgac gctggtcacg cgacctaggc tgggcttgta cacgtcagca 840  
 cgagcagggt cgtcaatgtt ctctgtgggac ctcatgctgc agatagtctc ctcaaggccg 900  
 ttgccctgtg ggctacgtgg gtgtctccac tcctcgctct cgtagggctg tcttagaggt 960  
 ggcttcacga cctggaaagg tcccttaaca cgaacgatgt ttctctgct gtcttggttg 1020  
 ttctgaagct gctgagccaa ctgaacgtca attttcaatg cttgagctat gacctgtgct 1080  
 tcgaacccgc tccacaagtt tttctgttct tgttgttgct gtgaaccgcc aaagcctccc 1140  
 tgctgggtgt ttccggccaa atggaacacg cggccgcaag ggcgaattct gcagatatcc 1200  
 atcacactgg cggccgctcg acgtaagctc aacaaatctt tagaaaatta attttatgtg 1260  
 acatatgcaa taatttgatt tggcaagata aactaataga ttttgcgatt tggagtttta 1320  
 aactctaaat aatctaaatc gttttcaatt ggtttaaata tatactctgc atttttaatc 1380  
 gtttttaatt aaaaaatata tatatatata tatatcttgc atttttaatc gttttcaatt 1440  
 taaaaaatat cttgcacgca gaacgctctc gagcggccgc ggatcctcag ggtcttttct 1500  
 tgcccacttt cttgaacgcc ggcaaaactca cgtttgttgt tcacggaagg ggtctaattg 1560  
 gaagagtat tccgggatgc gccgagacgt tcatggagtc accggtatct ggagaaggct 1620  
 aagggtcaggg tcagagtcaa gggttccgtg acatgcacca gaaagtagag cacctacggt 1680  
 gcggtgacac cattgcaaca ccactctgtg tagctcaatg gttctacaac aatggaaatg 1740  
 agcctctcat tcttggttgca gccgcggatc tcgccagcaa ccagaaccag cttgaccgca 1800  
 accttagacc atttttgata gccggaaca acccacaagg gcaggaatgg ctacaaggcc 1860  
 gaaagcaaca gaagcaaaaac aacatcttca atggcttcgc acctgagatc ttggctcaag 1920  
 ccttcaagat caatgtcgag acggctcagc agctccagaa ccagcaagat aaccgtggca 1980  
 acatcgtcaa ggtcaacgga ccttccggcg tcattaggcc acccttgaga cgcggcgaag 2040  
 gcggccaaca accacatgaa atagctaatt gtttagagga gactttgtgc accatgcgat 2100

```

gcactgaaaa cctc gatgac cegtcggatg ctgacgtgta caagccatca ctcggataca 2160
ttagcacact taacagctac aatcttccta tcctcagact tctccgcctt agcgcctctt 2220
gtggctccat ccgtaaaact cgaggtaagc tcaacaaatc tttagaaaat taattttatg 2280
tgacatatgc aataatttga tttggcaaga taaactaata gattttgcga tttggagttt 2340
taaactctaa ataactctaa tcgttttcaa ttgggtttaa tatatatctt gcatttttaa 2400
tcgtttttta ttaaaaaata tatatatata tatatatctt gcatttttaa tcgttttcaa 2460
tttaaaaaat atcttgacag cagaacgctg tcgagtttta cggatggagc cacgaagagc 2520
gctaaggcgg agaagtctga ggataggaag attgtagctg ttaagtgtgc taatgtatcc 2580
gagtgatggc ttgtacacgt cagcatccga cgggtcatcg aggttttcag tgcacgcacat 2640
ggtgcacaaa gtctcctcta aaccattagc tatttcattg ggttgttggc cgcttcgcc 2700
gcgtctcaag ggtggcctaa tgacgccgaa aggtccgttg accttgacga tgttgccacg 2760
gttatcttgc tgggttctgga gctgctgagc cgtctcgaca ttgatcttga aggttgagc 2820
caagatctca ggtgcgaagc cattgaagat gttgttttgc ttctgttgc ttcggccttg 2880
tagccattcc tgcccttgtg ggttgtttcc ggctatcaaa aatggctctaa ggttgcggtc 2940
aagctggttc tgggtgctgg cgagatccgc ggctgcaaca agaagagag gctcatttcc 3000
attgtttag aaccattgag ctacaccaga tgggtgttga atgggtgcac cgcaccgtag 3060
gtgctctact ttctggtgca tgtcacggaa cccttgactc tgaccctgac cttgaccttc 3120
tccaaatacc ggtgactcca tgaacgtctc ggcgcacccc ggaataactc ttcccattag 3180
accccttccg tgaacaacaa acgtgagttt gccggcgctt aagaaagtgg gcaagaaaag 3240
accctgagga tccgcggccg cgcctgcac taagctcagg taagctcaac aaatctttag 3300
aaaattaatt ttatgtgaca tatgcaataa tttgatttgg caagataaac taatagattt 3360
tgcgatttgg agtttttaaac tctaaataat ctaaactcgt ttcaattggg ttaaatatat 3420
atcttgcat tttaatcgtt tttaatataa aaatatatat atatatatat atcttgcat 3480
tttaatcgtt ttcaatttaa aaaatatctt gcacgcagaa cgctagggcc gcggccgagg 3540
atccatggct aacaagctct tcctcgtctg cgcaactctc gccctctgct tcctcctcac 3600
caacgcttcc atctaccgca ccgttgtcga attcgaagaa gatgacgcca gcaacccgt 3660
aggtccaaga cagagatgcc agaaggagtt tcagcaatca caacacctaa gagcttgcca 3720
gagatggatg agcaagcaaa tgaggcaagg acgtggtggt ggtccttccc tcgacgatga 3780
gttcgatttc gagggccccc agcagggata ccagctactc cagcagtgtc gcaacgagct 3840
tcgccaggaa gagccagttt gcgtttgcc caccttgaaa caagctgcca gggcagttag 3900
cctccaggga cagcacggac cattccaatc caggaaaatt taccagtcag ctaagtactt 3960
gcctaacatt tgcaagatcc agcaagttgg tgaatgtccc ttccagacca ccatcccttt 4020
cttccctcct tactactagg gtactcgagg taagctcaac aaatctttag aaaattaatt 4080
ttatgtgaca tatgcaataa tttgatttgg caagataaac taatagattt tgcgatttgg 4140
agtttttaaac tctaaataat ctaaactcgt ttcaattggg ttaaatatat atcttgcat 4200
tttaatcgtt tttaatataa aaatatatat atatatatat atcttgcat tttaatcgtt 4260
ttcaatttaa aaaatatctt gcacgcagaa cgctgtcgac taccctagta gtaaggaggg 4320
aagaaagga tgggtggtctg gaaggacat tcaccaactt gctggatctt gcaaatgtta 4380
ggcaagtact tagctgactg gtaaattttc ctggattgga atgggtccgtg ctgtccctgg 4440
aggctaactg ccctggcagc ttgtttcaag gtggggcaaa cgcaaactgg ctcttccctgg 4500
cgaagctcgt tgcagcactg ctggagtagc tggatccct gctgggggcc ctcgaaatcg 4560
aactcatcgt cgagggaagg accaccacca cgtccttgcc tcatttgctt gctcatccat 4620
ctctggcaag ctcttaggtg ttgtgattgc tgaaactcct tctggcatct ctgtcttgga 4680
cctacggggg tgctggcgct atcttcttgc aattcgacaa cgggtgcggta gatggaagcg 4740
ttggtgagga ggaagcagag ggcgagagtt gcgcagacga ggaagagctt gttagccatg 4800
gatcaattcg ccctatagtg agtcgtatta caattcactg gccgtcgttt tacaacgtcg 4860
tgactgggaa aaccctggcg ttacccaact taatcgctt gcagcacatc cccctttcgc 4920
cagctggcgt aatagcgaag aggcccgcac cgat 4954

```

&lt;210&gt; 84

&lt;211&gt; 4954

&lt;212&gt; RNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: RNA coding  
for dsRNA for suppression of multiple storage  
proteins

&lt;400&gt; 84

```

agcuugguac cgagcucgga uccacuagua acggccgcca gugugcugga auucgccccuu 60
gcggccgcggu guuccauuug gccggaaaca accagcaggg aggcuuuggc gguucacagc 120
aacaacaaga acagaaaaac uuguggagcg gguucgagcg acaggucaua gcucaagcau 180
ugaaaaauuga cguucaguug gcucagcagc uucagaacca acaagacagc agaggaaaca 240
ucguucgugu uaagggaccu uuccaggucg ugaggccacc ucuaagacag cccuacgaga 300
gcgaggagug gagacaccca cguagcccac agggcaacgg ccuugaggag acuaucugca 360
gcaugagguc ccacgagaac auugacgacc cugcucgugc ugacguguac aagcccagcc 420
uaggucgcu gaccagcguc aacagcuaua ccuugcccau cuuggaguau gucaggcuca 480
gugccacucg uggcgucuc caggguggau ccuucuguaa cauugacaa aacauguga 540
cacgucucc gucauauaga acuucaauu uuaauauguu uugcuaaaga aaaaaaaag 600
gaauaaaau cuaucaaaau cauuuuuuuaa acuuuuguuu acguucuaa auauuuuagg 660
auaugacuaa uuuuuuuuuu ugguaaaaau guuaauaucu auauuuuauu uauuuuagaa 720
aauguacuua caccucggag aacgccacga guggcacuga gccugacaua cuccaagau 780
ggcaaggua agcugugac gcuggucacg cgaccuaggc ugggcuugua cacgucagca 840
cgagcagggg cgucaauguu cucgugggac cucaugcugc agauagucuc cucaaggccg 900
uugcccuug ggcuacgug gugucuccac uccucgcucu cguagggcug ucuuagaggu 960
ggccucacga ccuggaaagg ucccuuaaca cgaacgaugu uuccucugcu gucuuugug 1020
uucugaagcu gcugagccaa cugaacguca auuuucaaug cuugagcuau gaccugugcg 1080
ucgaacccgc uccacaaguu uuucuguucu uguugugcu gugaaccgcc aaagccucc 1140
ugcugguugu uuccggccaa auggaacacg cggccgcaag ggcgaauucu gcagauaucc 1200
aucacacugg cggccgcucg acguaagcuc aacaaaucu uagaaaaua auuuuauug 1260
acauaugcaa uaauuugau uggcaagaa aacuaauaga uuugcgau uggaguuuua 1320
aacucuaaa auucuaaauc guuuucaau gguuuuaa uauaucuugc auuuuuuau 1380
guuuuuuau aaaaaaua uauauaua uauaucuugc auuuuuuau guuuucaau 1440
uaaaaaau cuugcacgca gaacgcuc cagcggccgc ggauccucag ggucuuuuc 1500
ugcccacuu cuugaacgcc ggcaaacuc cguuugugu ucacggaagg ggucuaaagg 1560
gaagaguau uccgggaug gccgagacgu ucauggaguc accgguauu ggagaagguc 1620
aaggucaggg ucagagucaa gguuuccgug acaugcacca gaaaguagag caccuacggu 1680
gcggugacac cauugcaaca ccaucuggug uagcucaaug guucuacaac aauggaaaug 1740
agccucuc auuuuugau gccgcggauc ucgccagcaa ccagaaccag cuugaccgca 1800
accuagacc auuuuugau gccggaaaca acccacaagg gcaggaaug cuacaaggcc 1860
gaaagcaaca gaagcaaac acaucuuca auggcucgc accugagauc ugggcucaag 1920
ccuucagau caaugucgag acggcucagc agcuccagaa ccagcaagau aaccguggca 1980
acaucgucaa ggucaacgga ccuucggcg ucauaggcc acccuugaga cgccggcgaag 2040
gcggccaaca acccaugaa auagcuauug guuuagagga gacuuugugc accaugcgau 2100
gcacugaaaa ccucgaugac ccgucggaug cugacgugua caagccaua cucggauaca 2160
uuagcacacu uaacagcuac aaucuuucca uccucagacu ucuccgccuu agcgcucuc 2220
guggcuccau ccguaaaacu cgagguaagc ucaacaauc uuuaaaaa auuuuuuau 2280
ugacauaugc aauauuuga uuuggcaaga uaaacuaa gauuuugcga uuuggaguuu 2340
uaaacucua auauucuaa ucguuuuca uugguuuua uauauaucu gcauuuuua 2400
ucguuuuuu uuaaaaaa uauauaua uauauaucu gcauuuuua ucguuuuca 2460
uuuaaaaa auucugcagc cagaacgcug ucgaguuuu cggauggagc cacgaagagc 2520
gcuaaggcgg agaagucuga ggauaggaag auugagcug uuaagugugc uaaugaucc 2580
gagugauggc uuguacacgu cagcauccga cgggucacug agguuuucag ugcaucgcau 2640
ggugcacaaa gucuccucua aaccuuagc uauuucaugu gguuugugc cgccuucgcc 2700
gcgucuaag gguggccua uagcgcgaa agguccguug accuugacga uguugccagc 2760
guauaucuugc ugguucugga gcugcugagc gcucgcaga uugaucuuga agguugagc 2820
caagaucuca ggugcgaagc cauugaagau guuuguuugc uucuguugcu uucggccuug 2880
uagccauucc ugcccuugug gguuguuucc ggcuaucuaa aauggucuaa gguugcgguc 2940
aagcugguuc ugguugcug cgagaucgc ggcugcaaca agaauagag gcucauuucc 3000
auuguugua aaccuuag cuacaccaga ugguguugca augguugac cgcaccguag 3060
gugcucua cuucuggugca ugucacgga cccuugacuc ugaccucgac cuugaccuuc 3120
uccaaauacc ggugacucca ugaacguc gcgcgaucc ggaauaacuc uucccauag 3180
acccuuccg ugaacaaca acgugaguuu gccggcguc aagaaagug gcaagaaaag 3240
accugagga uccgcggccg cgauagcau uagcucgag uaagcucaac aaucuuuag 3300
aaaauuau uuaugugaca uaugcauaa uuugauuug caagauaac uauuagauu 3360

```



```

ugcgauuugg aguuaaauaac ucuaaaauau cuaaaucguu uucaauuggu uuaaaauauau 3420
aucuugcauu uuuaaauucguu uuuaaauuaaa aaauauauau auauauauau aucuugcauu 3480
uuuaaauucguu uucaauuuuaa aaaaauaucuu gcacgcagaa cgcuaagggcc gcggccgcgg 3540
auccauggcu aacaagcucu uccucgucug cgcaacucuc gccucugcu uccuccucac 3600
caacgcuucc aucuaccgca ccguugucga auucgaagaa gaugacgcca gcaaccccg 3660
agguccaaga cagagaugcc agaaggaguu ucagcaauca caacaccuaa gagcuugcca 3720
gagauggaug agcaagcaaa ugagggaagg acgugguggu gguccuuccc ucgacgauga 3780
guucgauuuc gaggggcccc agcagggaaua ccagcuacuc cagcagugcu gcaacgagcu 3840
ucgccaggaa gagccaguuu gcguuugccc caccuugaaa caagcugcca gggcaguuag 3900
ccuccaggga cagcacggac cauucaaau caggaaaauu uaccagucag cuaaguacuu 3960
gccuaacauu ugcaagaucc agcaaguugg ugaauugucc uuccagacca ccauccuuu 4020
cuucccuccu uacuacuagg guacucgagg uaagcucaac aaauuuuag aaaaauaaau 4080
uuauugugaca uaugcaauaa uuugauuugg caagauaaac uaauagauuu ugcgauuugg 4140
aguuuuaaac ucuaaaauaa cuaaaucguu uucaauuggu uuaaaauauau aucuugcauu 4200
uuuaaauucguu uuuaaauuaaa aaauauauau auauauauau aucuugcauu uuuaaauucguu 4260
uucaauuuuaa aaaaauaucuu gcacgcagaa cgcugucgac uaccuagua guaaggagg 4320
aagaaaggga ugguggucug gaagggaau ucaccaacuu gcuggaucuu gcaaanguua 4380
ggcaaguacu uagcugacug guaaaauuuc cuggauugga augguccgug cucuccug 4440
aggcuaacug ccuggcagc uuguuucag guggggcaaa cgcaaacug cucuuccug 4500
cgaagcucgu ugcagcacug cuggaguagc ugguaucuu gcugggggccc cucgaaucg 4560
aacucaucgu cgagggaagg accaccacca cguccuugcc ucauuugcuu gcucauccau 4620
cucuggcaag cucuuaggug uugugauugc ugaacuccu ucuggcaucu cugucuugga 4680
ccuacggggg ugcuggcguc aucuucuuucg aaucgcagaa cggugcgua gauggaagcg 4740
uuggugagga ggaagcagag ggcgagaguu ggcgagacga ggaagagcuu guuagccaug 4800
gaucaauucg ccuauagug agucguauua caauucacug gccgucguuu uacaacgucg 4860
ugacugggaa aaccucggcg uuacccaacu uaucgccuu gcagcacauc cccuuucgc 4920
cagcuggcgu aaauagcgaag agggccgcac cgau 4954

```

&lt;210&gt; 85

&lt;211&gt; 4456

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: DNA  
construct coding for dsRNA for suppression of  
multiple storage proteins

&lt;400&gt; 85

```

agcttggtac cgagctcgga tccactagta acggccgcca gtgtgctgga attcgccctt 60
gcggccgcgg atcctcaggg tcttttcttg cccactttct tgaacgccgg caaactcacg 120
tttgttggtc acggaagggg tctaattgga agagtatttc cgggatgcgc cgagacgttc 180
atggagtcac cggtatattg agaaggtcaa ggtcagggtc agagtcaagg gttccgtgac 240
atgcaccaga aagtagagca cctacggtgc ggtgacacca ttgcaacacc atctggtgta 300
gctcaatggt tctacaacaa tggaaatgag cctctcattc ttgttgagc cgcggtatctc 360
gccagcaacc agaaccagct tgaccgcaac cttagaccat ttttgatagc cggaaacaac 420
ccacaagggc aggaatggct acaaggccga aagcaacaga agcaaaacaa catcttcaat 480
ggcttcgcac ctgagatctt ggctcaagcc ttcaagatca atgtcgagac ggctcagcag 540
ctccagaacc agcaagataa ccgtggcaac atcgtcaagg tcaacggacc tttcggcgctc 600
attaggccac cttgagacg cggcgaaggc ggccaacaac cacatgaaat agctaatggt 660
ttagaggaga ctttgtgcac catgcatgac actgaaaacc tcgatgaccc gtcggatgct 720
gacgtgtaca agccatcact cggatacatt agcacactta acagctacaa tcttcctatc 780
ctcagacttc tccgccttag cgtctctcgt ggctccatcc gtaaaaggat ccgcggccgc 840
aagggcgaat tctgcagatc cttcagggtc ttttcttgcc cactttcttg aacgccggca 900
aactcacgtt tgttggtcac ggaaggggtc taatgggaag agttattccg ggatgcgccg 960
agacgttcac ggagtcaccg gtatttgag aaggtcaagg tcagggtcag agtcaagggt 1020
tccgtgacat gcaccagaaa gtagagcacc tacggtgcgg tgacaccatt gcaacaccat 1080
ctggtgtagc tcaatggttc tacaacaatg gaaatgagcc tctcattctt gttgcagccg 1140
cggatctcgc cagcaaccag aaccagcttg accgcaacct tagaccattt ttgatagccg 1200

```

gaaacaaccc	acaagggcag	gaatggctac	aaggccgaaa	gcaacagaag	caaaacaaca	1260
tcttcaatgg	cttcgcacct	gagatcttgg	ctcaagcctt	caagatcaat	gtcgcagacgg	1320
ctcagcagct	ccagaaccag	caagataacc	gtggcaacat	cgtcaaggtc	aacggacctt	1380
tcggcgctcat	tagggcaccc	ttgagacgcg	gcgaaggcgg	ccaacaacca	catgaaatag	1440
ctaattggttt	agaggagact	ttgtgcacca	tgcgatgcac	tgaaaacctc	gatgacccgt	1500
cggatgctga	cgtgtacaag	ccatcactcg	gatacattag	cacacttaac	agctacaatc	1560
ttcctatcct	cagactttctc	cgccttagcg	ctcttcgtgg	ctccatccgt	aaaagatcct	1620
atggctaaca	agctcttctc	cgtctgcgca	actctcgccc	tctgcttctc	cctcaccaac	1680
gcttccatct	accgcaccgt	tgtcgaattc	gaagaagatg	acgccagcaa	ccccgtaggt	1740
ccaagacaga	gatgccagaa	ggagtttcag	caatcacaaac	acctaagagc	ttgccagaga	1800
tggatgagca	agcaaatgag	gcaaggacgt	ggtggtggtc	cttccctcga	cgatgagttc	1860
gatttcgagg	gccccagca	gggataccag	ctactccagc	agtgtctgcaa	cgagcttcgc	1920
caggaagagc	cagttttgcgt	ttgccccacc	ttgaaacaag	ctgccagggc	agttagcctc	1980
cagggacagc	acggaccatt	ccaatccagg	aaaattttacc	agtcagctaa	gtactttgcct	2040
aacattttgca	agatccagca	agttggtgaa	tgtcccttcc	agaccaccat	cccttttctc	2100
cctccttact	actagggttag	atatccatca	cactggcggc	cgctcgacgt	aagctcaaca	2160
aatcttttaga	aaattaattt	tatgtgacat	atgcaataat	ttgatttggc	aagataaact	2220
aatagattttt	gcgattttgga	gttttaaaact	ctaaataatc	taaatcgttt	tcaatttggtt	2280
taaatatata	tcttgcattt	ttaatcgttt	ttaattaaaa	aatatatata	tatatatata	2340
tcttgcattt	ttaatcgttt	tcaattttaa	aaatatcttg	cacgcagaac	gcctcgacta	2400
ccctagtagt	aaggagggaa	gaaagggatg	gtggtctgga	agggacattc	accaacttgc	2460
tggatcttgc	aaatgtttagg	caagtactta	gctgactggg	aaattttcct	ggatttggaat	2520
ggtccgtgct	gtccctggag	gctaactgcc	ctggcagctt	gtttcaaggt	ggggcaaacg	2580
caaactggct	cttcctggcg	aagctcgttg	cagcactgct	ggagtagctg	gtatccctgc	2640
tgggggccct	cgaaatcgaa	ctcatcgctc	agggaaaggac	caccaccacg	tccttgccctc	2700
atltgcttgc	tcatccatct	ctggcaagct	cttaggtgtt	gtgattgctg	aaactccttc	2760
tggcatctct	gtcttggacc	tacgggggtt	ctggcgctcat	cttcttcgaa	ttcgacaacg	2820
gtgcggtaga	tggaaagcgtt	ggtgaggagg	aagcagaggg	cgagagttgc	gcagacgagg	2880
aagagcttgt	tagccatagg	atcttttacg	gatggagcca	cgaagagcgc	taaggcggag	2940
aagtctgagc	ataggaagat	tgtagctggt	aagtgtgcta	atgtatccga	gtgatggctt	3000
gtacacgtga	gcataccgacg	ggtcatcgag	gttttcagtg	catcgcatgg	tgcacaaagt	3060
ctcctctaaa	ccattagcta	tttcatgtgg	ttgttggcgg	ccttcgcgcg	gtctcaaggg	3120
tggcctaattg	acgccgaaag	gtccgttgac	cttgacgatg	ttgccacggt	tatcttgcgtg	3180
gttctggagc	tgtctgagccg	tctcgacatt	gatcttgaag	gcttgagcca	agatctcagg	3240
tgcgaagcca	ttgaagatgt	tgttttgcct	ctgttgcctt	cggccttgta	gccattcctg	3300
cccttgtggg	ttgtttccgg	ctatcaaaaa	tggctctaagg	ttgcggtcaa	gctgggttctg	3360
gttgctggcg	agatccgcgg	ctgcaacaag	aatgagaggg	tcatttccat	tgttgtagaa	3420
ccattgagct	acaccagatg	gtgttgcaat	ggtgtcaccg	caccgtaggt	gctctacttt	3480
ctggtgcatg	tcacggaacc	cttgactctg	accctgacct	tgaccttctc	caaataaccg	3540
tgactccatg	aacgtctcgg	cgcatacccg	aataactctt	cccattagac	cccttcctg	3600
aacaacaaac	gtgagtttgc	cggcggttcaa	gaaagtgggc	aagaaaagac	cctgaaggat	3660
ctgcagaatt	cgcccttgcg	gccgcggatc	cttttacgga	tggagccacg	aagagcgcta	3720
aggcgagaa	gtctgaggat	aggaagattg	tagctgttaa	gtgtgcta	gtatccgagt	3780
gatggcttgt	acacgtcagc	atccgacggg	tcatcgaggt	tttcagtgca	tcgcagtgtg	3840
cacaaagtct	cctctaaacc	attagctatt	tcatgtggtt	gttggccgcc	ttcgccgcgt	3900
ctcaagggtg	gcctaattgac	gccgaaagg	ccgttgacct	tgacgatgtt	gccacgggta	3960
tcttgctggg	tctggagctg	ctgagccgtc	tcgacattga	tcttgaaggc	ttgagccaag	4020
atctcaggtg	cgaagccatt	gaagatgttg	ttttgcttct	gttgccttctg	gccttgttagc	4080
cattcctgcc	cttgtgggtt	gtttccggct	atcaaaaatg	gtctaagggt	gcggtcaagc	4140
tggttctggg	tgttgccgag	atccgcggct	gcaacaagaa	tgagaggctc	atttccattg	4200
ttgtagaacc	attgagctac	accagatggg	gttgcaatgg	tgtcacccga	ccgtaggtgc	4260
tctactttct	ggtgcatgtc	acggaaccct	tgactctgac	cctgaccttg	accttctcca	4320
aataccggtg	actccatgaa	cgtctcggcg	catcccgga	taactcttcc	cattagaccc	4380
cttccgtgaa	caacaaacgt	gagtttgccg	gcgttcaaga	aagtgggcaa	gaaaagaccc	4440
tgactcgagc	atgcat					4456

&lt;210&gt; 86

&lt;211&gt; 4456

&lt;212&gt; RNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: RNA coding  
for dsRNA for suppression of multiple storage  
proteins

&lt;400&gt; 86

```

agcuugguac cgagcucgga uccacuagua acggccgcca gugugcugga auucgccccu 60
gcgggccgagg auccucaggg ucuuuuucug cccacuuucu ugaacgcccgg caaacucacg 120
uuuuguuguuc acggaagggg ucuaauuggg agaguuauuc cgggaugcgc cgagacguuc 180
auggagucac cggauuuugg agaaggucac ggucaggguc agagucaagg guuccgugac 240
augcaccaga aaguagagca ccuacggugc ggugacacca uugcaacacc aucuggugua 300
gcucaauggu ucuaacaaca uggaaugag ccucucauuc uuguugcagc cgcggaucuc 360
gccagcaacc agaaccagcu ugaccgcaac cuuagaccuau uuuugauagc cggaaacaac 420
ccacaagggc aggaauaggc acaaggccga aagcaacaga agcaaaaca caucucaau 480
ggcuucgcac cugagaucuu ggcucaagcc uucaagauc augucgagac ggcucagcag 540
cuccagaacc agcaagauaa ccguggcaac aucgucaagg ucaacggacc uuucggcguc 600
auuaggccac ccuugagacg cggcgaaggc ggccaacaac cacaugaaau agcuauuggu 660
uuagaggaga cuuugugcac caugcgauag acugaaaacc ucgaugaccc gucggaugcu 720
gacguguaca agccaucacu cggauacauu agcacacuua acagcuacaa ucuuccuau 780
cucagacuuc uccgccuua ggcucuucgu ggcuccaucc guaaaaggau ccgcgccgc 840
aagggcgaau ucugcagauc cuucaggguc uuuuucugcc cacuuucug aacgccggca 900
aacucacguu uguuguucac ggaagggguc uauugggaag aguauuuccg ggaugcgccg 960
agacguucau ggagucaccg guauuuggag aaggucacag ucagggucag agucaagggu 1020
uccgugacau gcaccagaaa guagagcacc uacggugcgg ugacaccuau gcaacaccu 1080
cugguguauc ucaauuggu uacaacaau gaaugagcc ucucauucuu guugcagccg 1140
cggauucgc cagcaaccag aaccagcuug accgcaaccu uagaccuuu uugauagccg 1200
gaaacaacc acaagggcag gaauggcuac aaggccgaaa gcaacagaag caaaacaaca 1260
ucuucaaug cuucgcaccu gagaucuug cucaagccuu caagaucuu gucgagacg 1320
cucagcagcu ccagaaccag caagauaacc guggcaacu cgucaagguc aacggaccu 1380
ucggcgucuu uaggccacc uugagcggc gcgaaggcgg ccaacaacca caugaaauag 1440
cuauuguuu agaggagacu uugugcacca ugcgauagc ugaauaccu gaugacccgu 1500
cggaugcuga cguguacaag ccaucacucg gauacauuag cacacuuaac agcuacaau 1560
uuccuauccu cagacuucuc cgccuagcgg cucuucgugg cuccaucgu aaaagauc 1620
auggcuaaca agcucuuccu cgucugcgca acucucgccc ucugcuuccu ccucaccaac 1680
gcuuccauuc accgcaccgu ugucgaauc gaagaagau acgccagcaa ccccguaag 1740
ccaagacaga gaugccagaa ggaguucag caucacaac accuaagagc uugccagaga 1800
uggaugagca agcaauagag gcaaggacgu gguggugguc cuuccucga cgaugaguuc 1860
gauuucgagg gccccagca gggauaccag cuacuccagc agugcugcaa cgagcuucg 1920
caggaagagc caguugcggu uugccccacc uugaaacaag cugccagggc aguuagccu 1980
cagggacagc acggaccuau ccauuccagg aaauuuuacc agucagcuua guacuugccu 2040
aacauuugca agauccagca aguuggugaa ugucuccuucc agaccaccu cccuucuc 2100
ccuccuacu acuagguag auauccauc cacuggcgcc cgucgcagcu aagcuaca 2160
aaucuuuaga aaauuaauu uauugacau augcauaau ugaauuuggc aagaauaac 2220
aaauaguuu gcauuugga guuuuaaau cuauuaau uaaucguuu ucaauuggu 2280
uaauauaua ucuugcauu uaaucguuu uaaauuaaa auauauaua uauauaua 2340
ucugcauuu uaaucguuu ucauuuaaa aaauaucuug cagcgagaa gccucgacua 2400
ccuaguagu aaggaggga gaaagggaug guggucugga agggacauuc accacuugc 2460
uggaucuuuc aaauuuagg caaguacua gcugacuggu aaauuuuccu ggauuggaau 2520
gguccgugcu guccucggag gcuaacugcc cuggcagcuu guuucaggu ggggcaaacg 2580
caaacuggcu cuuccuggcg aagcucguug cagcacugcu ggaguagcu guauccugc 2640
ugggggccc cgaauucgaa cucaucguc aggaaggac caccaccag uccuuccuc 2700
auuugcuugc ucauccauc cuggcaagcu cuuagguu gugauugcu aaacucuc 2760
uggcaucuc gucuuggacc uacggguug cuggcguc cuucucgaa uucgacaacg 2820
gugcgguaga uggaagcguu ggugaggag aagcagagg cgagaguug gcagacgag 2880
aagagcuugu uagccauagg aucuuuacg gauggagcca cgaagagcg uaaggcggag 2940

```

```

aagucugagg auaggaagau uguagcuguu aagugugcua auguauccga gugauggcuu 3000
guacacguca gcauccgacg ggucaucgag guuuucagug caucgcaugg ugcacaaagu 3060
cuccucuaaa ccauagcua uucaugugg uuguuggccg ccuucgccc gucucaaggg 3120
uggccuaaag acgccgaaag guccguugac cuugacgaug uugccacggg uaucuugcug 3180
guucuggagc ugcugagccg ucucgacauu gaucugaag gcuugagcca agaucucagg 3240
ugcgaagcca uugaagaugu uguuuugcuu cuguugcuiu cggccuugua gccauuccug 3300
cccuuguggg uuguuuccgg cuaucaaaaa uggucuaagg uugcggucua gcugguucug 3360
guugcuggcg agauccgcgg cugcaacaag aaugagaggc ucauuuccau uguuguagaa 3420
ccaugagcu acaccagau guguugcau ggugucaccg caccguaggu gcucuacuuu 3480
cuggugcaug ucacggaacc cuugacucug acccugaccu ugaccuucuc caaaauaccgg 3540
ugacuccaug aacgucucgg cgcaucccg gaaagugggc aagaaaagac ccugaaggau 3600
aacaacaaac gugaguuugc cggcguucaa gaaagugggc aagaaaagac ccugaaggau 3660
cugcagaauu cgcccuugcg gccgcggau cuuuuacgga uggagccacg aagagcgcu 3720
aggcggagaa gucugaggau aggaagauug uagcuguuaa gugugcuaau guauccgagu 3780
gauggcuguu acacgucagc auccgacggg ucaucgaggu uuucagugca ucgcauggug 3840
cacaagucu ccucuaaacc auuagcuau ucaugugguu guuggccgcc uucgcccgcu 3900
cucaagggug gccuaaugac gccgaaaggu ccguugaccu ugacgauguu gccacgguaa 3960
ucuugcuggu ucuggagcug cugagccguc ucgacauuga ucuugaaggc uugagccaag 4020
aucucaggug cgaagccau gaagauguug uuugcuucu guugcuuucg gccuuguagc 4080
cauuccugcc cuuguggguu guuuccggcu aucaaaaaug gucuaagggu gccgucaagc 4140
ugguucuggu ugcuggcgag auccgcggcu gcaacaagaa ugagaggcuc auuuccauug 4200
uuguagaacc auugagcuac accagauggu guugcaaug ugucaccgca ccguaggugc 4260
ucuacuuucu ggugcaugc acggaaccu ugacucugac ccugaccuug accuucucca 4320
aaauaccggug acuccaugaa cgucucggcg caucccgaa uaacucuucc cauugaccc 4380
cuuccgugaa caacaaacgu gaguuugccg gcguucaaga aagugggcaa gaaaagacc 4440
ugacucgagc augcau

```

4456

&lt;210&gt; 87

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

&lt;400&gt; 87

aaaaggcctg tggtccattt ggccggaaac aac

33

&lt;210&gt; 88

&lt;211&gt; 31

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

&lt;400&gt; 88

aaagatatca ccctggagaa cgccacgagt g

31

&lt;210&gt; 89

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz:  
oligonucleotide primer

&lt;400&gt; 89

aaaaggccta tggctaacaa gctcttcctc gtc

33

<210> 90  
 <211> 32  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
           oligonucleotide primer  
 <400> 90  
 aaagatatcc tagtagtaag gagggaagaa ag 32  
 <210> 91  
 <211> 32  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
           oligonucleotide primer  
 <400> 91  
 ccgctcgagc tcagggctctt ttcttgccca ct 32  
 <210> 92  
 <211> 32  
 <212> DNA  
 <213> Künstliche Sequenz  
 <220>  
 <223> Beschreibung der künstlichen Sequenz:  
           oligonucleotide primer  
 <400> 92  
 ccggtcgacc tagtagtaag gagggaagaa ag 32  
 <210> 93  
 <211> 1500  
 <212> DNA  
 <213> Arabidopsis thaliana  
 <220>  
 <221> CDS  
 <222> (1)..(1497)  
 <223> globuline-like protein  
 <400> 93  
 atg act aga ttt gcg gta ttg cca ctc tct gtt ctt ctt ctc gtt ctc 48  
 Met Thr Arg Phe Ala Val Leu Pro Leu Ser Val Leu Leu Leu Val Leu  
   1                  5                  10                  15  
 ttg ttc ctc tgc act gag tcg ttg gct aag tcg gag gag tct gaa gaa 96  
 Leu Phe Leu Cys Thr Glu Ser Leu Ala Lys Ser Glu Glu Ser Glu Glu  
                   20                  25                  30  
 tac gac gtc gct gtg cct tca tgt tgc ggg ttt tcg tct cct ctt ctg 144  
 Tyr Asp Val Ala Val Pro Ser Cys Cys Gly Phe Ser Ser Pro Leu Leu  
           35                  40                  45  
 att aag aaa gat caa tgg aaa cca atc ttc gag acg aag ttt gga cag 192  
 Ile Lys Lys Asp Gln Trp Lys Pro Ile Phe Glu Thr Lys Phe Gly Gln  
   50                  55                  60  
 atc tca acc gtt caa atc ggc aat gga tgc ggt gga atg gga cct tac 240  
 Ile Ser Thr Val Gln Ile Gly Asn Gly Cys Gly Gly Met Gly Pro Tyr  
   65                  70                  75                  80

95

aag	ata	cat	tcc	ata	acg	ttg	gag	cca	aac	aca	att	ttg	ctc	cct	ctt	288
Lys	Ile	His	Ser	Ile	Thr	Leu	Glu	Pro	Asn	Thr	Ile	Leu	Leu	Pro	Leu	
			85						90					95		
ctt	ctt	cat	tcc	gac	atg	gtc	ttc	ttc	gtt	gac	tct	gga	agt	ggg	att	336
Leu	Leu	His	Ser	Asp	Met	Val	Phe	Phe	Val	Asp	Ser	Gly	Ser	Gly	Ile	
			100				105					110				
ctg	aat	tgg	gtt	gac	gag	gaa	gcg	aag	agt	acc	gag	atc	aga	cta	gga	384
Leu	Asn	Trp	Val	Asp	Glu	Glu	Ala	Lys	Ser	Thr	Glu	Ile	Arg	Leu	Gly	
		115					120					125				
gac	gtt	tac	agg	cta	cgt	ccc	ggg	tcg	gtt	ttc	tac	tta	cag	agc	aaa	432
Asp	Val	Tyr	Arg	Leu	Arg	Pro	Gly	Ser	Val	Phe	Tyr	Leu	Gln	Ser	Lys	
	130					135					140					
ccg	gat	cct	tgc	ttt	ggg	gcc	tat	tcg	agt	atc	aca	gat	cta	atg	ttt	480
Pro	Asp	Pro	Cys	Phe	Gly	Ala	Tyr	Ser	Ser	Ile	Thr	Asp	Leu	Met	Phe	
	145				150					155					160	
ggg	ttt	gat	gag	aca	att	ctc	cag	tca	gct	ttt	ggg	gtt	cct	gag	ggg	528
Gly	Phe	Asp	Glu	Thr	Ile	Leu	Gln	Ser	Ala	Phe	Gly	Val	Pro	Glu	Gly	
			165					170					175			
att	ata	gag	ctg	atg	agg	aac	cgt	acg	aag	cca	cca	ttg	atc	gtg	agt	576
Ile	Ile	Glu	Leu	Met	Arg	Asn	Arg	Thr	Lys	Pro	Pro	Leu	Ile	Val	Ser	
			180					185					190			
gag	acg	ctt	tgc	aca	cca	ggg	gtg	gcc	aac	act	tgg	cag	ctc	caa	ccg	624
Glu	Thr	Leu	Cys	Thr	Pro	Gly	Val	Ala	Asn	Thr	Trp	Gln	Leu	Gln	Pro	
		195					200					205				
cga	tta	cta	aag	ctc	ttt	gct	gga	agt	gca	gac	ttg	gtg	gat	aac	aag	672
Arg	Leu	Leu	Lys	Leu	Phe	Ala	Gly	Ser	Ala	Asp	Leu	Val	Asp	Asn	Lys	
	210				215						220					
aag	aag	aaa	gag	aag	aaa	gag	aag	aag	gag	aag	gtg	aag	aaa	gcg	aag	720
Lys	Lys	Lys	Glu	Lys	Lys	Glu	Lys	Lys	Glu	Lys	Val	Lys	Lys	Ala	Lys	
	225				230				235					240		
aca	ttc	aat	gtt	ttc	gaa	tct	gaa	ccg	gat	ttc	gag	agc	ccc	tac	ggg	768
Thr	Phe	Asn	Val	Phe	Glu	Ser	Glu	Pro	Asp	Phe	Glu	Ser	Pro	Tyr	Gly	
			245					250					255			
cgt	act	ata	acg	att	aac	agg	aag	gat	ctg	aaa	gtt	tta	aaa	ggc	tca	816
Arg	Thr	Ile	Thr	Ile	Asn	Arg	Lys	Asp	Leu	Lys	Val	Leu	Lys	Gly	Ser	
		260						265				270				
atg	gtt	gga	gtt	tcc	atg	gtg	aat	ctc	act	caa	gga	tcg	atg	atg	gga	864
Met	Val	Gly	Val	Ser	Met	Val	Asn	Leu	Thr	Gln	Gly	Ser	Met	Met	Gly	
		275					280					285				
ccg	cac	tgg	aat	cca	tgg	gct	tgc	gag	att	tcg	att	gta	ttg	aaa	gga	912
Pro	His	Trp	Asn	Pro	Trp	Ala	Cys	Glu	Ile	Ser	Ile	Val	Leu	Lys	Gly	
		290				295						300				
gca	gga	atg	gtt	agg	gtt	ctt	agg	tct	tcg	ata	tca	tca	aac	aca	tca	960
Ala	Gly	Met	Val	Arg	Val	Leu	Arg	Ser	Ser	Ile	Ser	Ser	Asn	Thr	Ser	
	305				310					315					320	
tca	gag	tgt	aag	aac	gtg	agg	ttt	aaa	gta	gag	gaa	gga	gat	att	ttc	1008
Ser	Glu	Cys	Lys	Asn	Val	Arg	Phe	Lys	Val	Glu	Glu	Gly	Asp	Ile	Phe	
			325					330					335			
gca	gtt	cca	cgg	tta	cat	cca	atg	gct	caa	atg	tct	ttc	aac	aat	gac	1056
Ala	Val	Pro	Arg	Leu	His	Pro	Met	Ala	Gln	Met	Ser	Phe	Asn	Asn	Asp	
			340					345					350			

## 96

tcg tta gtg ttc gtt ggg ttt act act tca gct aag aat aac gag ccg	1104
Ser Leu Val Phe Val Gly Phe Thr Thr Ser Ala Lys Asn Asn Glu Pro	
355 360 365	
cag ttc tta gcc ggg gag gac tcg gct ttg cgg atg ctt gac cgg caa	1152
Gln Phe Leu Ala Gly Glu Asp Ser Ala Leu Arg Met Leu Asp Arg Gln	
370 375 380	
gta ttg gct gca tcg ctt aat gtg agt agt gtg acg att gat gga ttg	1200
Val Leu Ala Ala Ser Leu Asn Val Ser Ser Val Thr Ile Asp Gly Leu	
385 390 395 400	
ttg gga gct cag aag gaa gct gtt atc ttg gaa tgt cat tct tgt gcg	1248
Leu Gly Ala Gln Lys Glu Ala Val Ile Leu Glu Cys His Ser Cys Ala	
405 410 415	
gaa gga gag ata gag aag ctt aag gtg gag ata gag agg aag aaa ata	1296
Glu Gly Glu Ile Glu Lys Leu Lys Val Glu Ile Glu Arg Lys Lys Ile	
420 425 430	
gat gat gag agg aag agg aga cat gat gaa agg aag aaa gaa gaa gaa	1344
Asp Asp Glu Arg Lys Arg Arg His Asp Glu Arg Lys Lys Glu Glu Glu	
435 440 445	
gag gcg aag aga gaa gag gaa gag agg agg aaa cgt gaa gaa gaa gaa	1392
Glu Ala Lys Arg Glu Glu Glu Glu Arg Arg Lys Arg Glu Glu Glu Glu	
450 455 460	
gag aag aag cgg tgg cca cct caa caa cca cca caa gaa gaa gaa ctt	1440
Glu Lys Lys Arg Trp Pro Pro Gln Gln Pro Pro Gln Glu Glu Glu Leu	
465 470 475 480	
agg gaa cgg caa cta ccg atg gag aaa gaa tgg gaa atg gaa ggt gaa	1488
Arg Glu Arg Gln Leu Pro Met Glu Lys Glu Trp Glu Met Glu Gly Glu	
485 490 495	
gag gag agt taa	1500
Glu Glu Ser	
<210> 94	
<211> 499	
<212> PRT	
<213> Arabidopsis thaliana	
<400> 94	
Met Thr Arg Phe Ala Val Leu Pro Leu Ser Val Leu Leu Leu Val Leu	
1 5 10 15	
Leu Phe Leu Cys Thr Glu Ser Leu Ala Lys Ser Glu Glu Ser Glu Glu	
20 25 30	
Tyr Asp Val Ala Val Pro Ser Cys Cys Gly Phe Ser Ser Pro Leu Leu	
35 40 45	
Ile Lys Lys Asp Gln Trp Lys Pro Ile Phe Glu Thr Lys Phe Gly Gln	
50 55 60	
Ile Ser Thr Val Gln Ile Gly Asn Gly Cys Gly Gly Met Gly Pro Tyr	
65 70 75 80	
Lys Ile His Ser Ile Thr Leu Glu Pro Asn Thr Ile Leu Leu Pro Leu	
85 90 95	
Leu Leu His Ser Asp Met Val Phe Phe Val Asp Ser Gly Ser Gly Ile	
100 105 110	
Leu Asn Trp Val Asp Glu Glu Ala Lys Ser Thr Glu Ile Arg Leu Gly	
115 120 125	

Asp	Val	Tyr	Arg	Leu	Arg	Pro	Gly	Ser	Val	Phe	Tyr	Leu	Gln	Ser	Lys	130	135	140
Pro	Asp	Pro	Cys	Phe	Gly	Ala	Tyr	Ser	Ser	Ile	Thr	Asp	Leu	Met	Phe	145	150	155
Gly	Phe	Asp	Glu	Thr	Ile	Leu	Gln	Ser	Ala	Phe	Gly	Val	Pro	Glu	Gly	165	170	175
Ile	Ile	Glu	Leu	Met	Arg	Asn	Arg	Thr	Lys	Pro	Pro	Leu	Ile	Val	Ser	180	185	190
Glu	Thr	Leu	Cys	Thr	Pro	Gly	Val	Ala	Asn	Thr	Trp	Gln	Leu	Gln	Pro	195	200	205
Arg	Leu	Leu	Lys	Leu	Phe	Ala	Gly	Ser	Ala	Asp	Leu	Val	Asp	Asn	Lys	210	215	220
Lys	Lys	Lys	Glu	Lys	Lys	Glu	Lys	Lys	Glu	Lys	Val	Lys	Lys	Ala	Lys	225	230	235
Thr	Phe	Asn	Val	Phe	Glu	Ser	Glu	Pro	Asp	Phe	Glu	Ser	Pro	Tyr	Gly	245	250	255
Arg	Thr	Ile	Thr	Ile	Asn	Arg	Lys	Asp	Leu	Lys	Val	Leu	Lys	Gly	Ser	260	265	270
Met	Val	Gly	Val	Ser	Met	Val	Asn	Leu	Thr	Gln	Gly	Ser	Met	Met	Gly	275	280	285
Pro	His	Trp	Asn	Pro	Trp	Ala	Cys	Glu	Ile	Ser	Ile	Val	Leu	Lys	Gly	290	295	300
Ala	Gly	Met	Val	Arg	Val	Leu	Arg	Ser	Ser	Ile	Ser	Ser	Asn	Thr	Ser	305	310	315
Ser	Glu	Cys	Lys	Asn	Val	Arg	Phe	Lys	Val	Glu	Glu	Gly	Asp	Ile	Phe	325	330	335
Ala	Val	Pro	Arg	Leu	His	Pro	Met	Ala	Gln	Met	Ser	Phe	Asn	Asn	Asp	340	345	350
Ser	Leu	Val	Phe	Val	Gly	Phe	Thr	Thr	Ser	Ala	Lys	Asn	Asn	Glu	Pro	355	360	365
Gln	Phe	Leu	Ala	Gly	Glu	Asp	Ser	Ala	Leu	Arg	Met	Leu	Asp	Arg	Gln	370	375	380
Val	Leu	Ala	Ala	Ser	Leu	Asn	Val	Ser	Ser	Val	Thr	Ile	Asp	Gly	Leu	385	390	395
Leu	Gly	Ala	Gln	Lys	Glu	Ala	Val	Ile	Leu	Glu	Cys	His	Ser	Cys	Ala	405	410	415
Glu	Gly	Glu	Ile	Glu	Lys	Leu	Lys	Val	Glu	Ile	Glu	Arg	Lys	Lys	Ile	420	425	430
Asp	Asp	Glu	Arg	Lys	Arg	Arg	His	Asp	Glu	Arg	Lys	Lys	Glu	Glu	Glu	435	440	445
Glu	Ala	Lys	Arg	Glu	Glu	Glu	Glu	Arg	Arg	Lys	Arg	Glu	Glu	Glu	Glu	450	455	460
Glu	Lys	Lys	Arg	Trp	Pro	Pro	Gln	Gln	Pro	Pro	Gln	Glu	Glu	Glu	Leu	465	470	475
Arg	Glu	Arg	Gln	Leu	Pro	Met	Glu	Lys	Glu	Trp	Glu	Met	Glu	Gly	Glu	485	490	495
Glu	Glu	Ser																



<210> 95  
 <211> 1284  
 <212> DNA  
 <213> Glycine max  
 <220>  
 <221> CDS  
 <222> (1)..(1281)  
 <223> 7S seed globuline

<400> 95  
 atg gct tcc atc ctc cac tac ttt tta gcc ctc tct ctt tct tgc tct 48  
 Met Ala Ser Ile Leu His Tyr Phe Leu Ala Leu Ser Leu Ser Cys Ser  
 1 5 10 15  
 ttt ctt ttc ttc tta tcc gac tca gtc acc cct aca aaa cca ata aac 96  
 Phe Leu Phe Phe Leu Ser Asp Ser Val Thr Pro Thr Lys Pro Ile Asn  
 20 25 30  
 ctt gtt gtt cta ccc gtt caa aat gat ggt tcc aca ggg ctc cat tcg 144  
 Leu Val Val Leu Pro Val Gln Asn Asp Gly Ser Thr Gly Leu His Ser  
 35 40 45  
 gcc aac ctc caa aaa aga acc cct cta atg caa gta cca gtc ctg gtg 192  
 Ala Asn Leu Gln Lys Arg Thr Pro Leu Met Gln Val Pro Val Leu Val  
 50 55 60  
 gac ctc aat gga aat cac ttg tgg gtt aac tgt gag cag cag tac tca 240  
 Asp Leu Asn Gly Asn His Leu Trp Val Asn Cys Glu Gln Gln Tyr Ser  
 65 70 75 80  
 tcc aaa acg tac caa gca ccc ttc tgc cac tcc acc caa tgc tct aga 288  
 Ser Lys Thr Tyr Gln Ala Pro Phe Cys His Ser Thr Gln Cys Ser Arg  
 85 90 95  
 gcc aac acc cac caa tgc ctc agt tgc ccc gcg gca tca agg cca ggg 336  
 Ala Asn Thr His Gln Cys Leu Ser Cys Pro Ala Ala Ser Arg Pro Gly  
 100 105 110  
 tgc cac aaa aac acg tgt ggc ctc atg tcc act aat ccc atc acc caa 384  
 Cys His Lys Asn Thr Cys Gly Leu Met Ser Thr Asn Pro Ile Thr Gln  
 115 120 125  
 caa acc ggt tta ggt gaa cta gga gaa gac gtt ctt gca atc cac gcc 432  
 Gln Thr Gly Leu Gly Glu Leu Gly Glu Asp Val Leu Ala Ile His Ala  
 130 135 140  
 aca caa ggg tcg acc caa caa ctt ggc cca ttg gtc aca gtc cca caa 480  
 Thr Gln Gly Ser Thr Gln Gln Leu Gly Pro Leu Val Thr Val Pro Gln  
 145 150 155 160  
 ttc ctc ttt tct tgt gca cct tcc ttc ctt gtt caa aag ggt ctt cct 528  
 Phe Leu Phe Ser Cys Ala Pro Ser Phe Leu Val Gln Lys Gly Leu Pro  
 165 170 175  
 aga aac act caa ggt gtg gct ggg tta ggc cat gca cca att tct ctt 576  
 Arg Asn Thr Gln Gly Val Ala Gly Leu Gly His Ala Pro Ile Ser Leu  
 180 185 190  
 cca aac caa ctc gct tcc cac ttt ggc cta caa cgc caa ttc acc act 624  
 Pro Asn Gln Leu Ala Ser His Phe Gly Leu Gln Arg Gln Phe Thr Thr  
 195 200 205  
 tgc ctt tct cgc tac cct act tca aag ggt gct ata ata ttc ggg gat 672  
 Cys Leu Ser Arg Tyr Pro Thr Ser Lys Gly Ala Ile Ile Phe Gly Asp  
 210 215 220

## 99

gca cct aac aac atg cga cag ttt caa aac caa gat att ttc cac gat 720  
 Ala Pro Asn Asn Met Arg Gln Phe Gln Asn Gln Asp Ile Phe His Asp  
 225 230 235 240  
 ttg gcc ttc acc cca tta acc atc acc ctg cag gga gag tac aac gtg 768  
 Leu Ala Phe Thr Pro Leu Thr Ile Thr Leu Gln Gly Glu Tyr Asn Val  
 245 250 255  
 aga gtc aac tca ata aga atc aac cag cac agt gtg ttc cca ctg aac 816  
 Arg Val Asn Ser Ile Arg Ile Asn Gln His Ser Val Phe Pro Leu Asn  
 260 265 270  
 aag ata tca tcc acc atc gta ggg tcg acc tct gga gga acc atg att 864  
 Lys Ile Ser Ser Thr Ile Val Gly Ser Thr Ser Gly Gly Thr Met Ile  
 275 280 285  
 agc acc tca act cct cac atg gtt ctc cag caa tcc gtg tac cag gct 912  
 Ser Thr Ser Thr Pro His Met Val Leu Gln Gln Ser Val Tyr Gln Ala  
 290 295 300  
 ttc act cag gtg ttt gct cag cag cta cca aag caa gca cag gtg aaa 960  
 Phe Thr Gln Val Phe Ala Gln Gln Leu Pro Lys Gln Ala Gln Val Lys  
 305 310 315 320  
 tct gtg gca cca ttt ggg tta tgc ttc aac tcc aac aag atc aat gca 1008  
 Ser Val Ala Pro Phe Gly Leu Cys Phe Asn Ser Asn Lys Ile Asn Ala  
 325 330 335  
 tat cct agc gtg gac ctt gtg atg gac aag ccc aat ggt cct gtt tgg 1056  
 Tyr Pro Ser Val Asp Leu Val Met Asp Lys Pro Asn Gly Pro Val Trp  
 340 345 350  
 aga atc tct ggt gag gac ttg atg gtg cag gca caa cct ggg gtc acg 1104  
 Arg Ile Ser Gly Glu Asp Leu Met Val Gln Ala Gln Pro Gly Val Thr  
 355 360 365  
 tgt ttg ggt gtt atg aat gga gga atg caa cct aga gct gaa att acc 1152  
 Cys Leu Gly Val Met Asn Gly Gly Met Gln Pro Arg Ala Glu Ile Thr  
 370 375 380  
 tta ggg gca cgt cag ttg gaa gag aac ctg gtg gtg ttc gat ctt gca 1200  
 Leu Gly Ala Arg Gln Leu Glu Glu Asn Leu Val Val Phe Asp Leu Ala  
 385 390 395 400  
 agg tca agg gtt ggg ttt agc acc tca tca ctg cac tcg cat gga gtc 1248  
 Arg Ser Arg Val Gly Phe Ser Thr Ser Ser Leu His Ser His Gly Val  
 405 410 415  
 aaa tgt gct gac ctc ttc aac ttt gcc aat gca tga 1284  
 Lys Cys Ala Asp Leu Phe Asn Phe Ala Asn Ala  
 420 425

&lt;210&gt; 96

&lt;211&gt; 427

&lt;212&gt; PRT

&lt;213&gt; Glycine max

&lt;400&gt; 96

Met Ala Ser Ile Leu His Tyr Phe Leu Ala Leu Ser Leu Ser Cys Ser  
 1 5 10 15  
 Phe Leu Phe Phe Leu Ser Asp Ser Val Thr Pro Thr Lys Pro Ile Asn  
 20 25 30  
 Leu Val Val Leu Pro Val Gln Asn Asp Gly Ser Thr Gly Leu His Ser  
 35 40 45

## 100

Ala	Asn	Leu	Gln	Lys	Arg	Thr	Pro	Leu	Met	Gln	Val	Pro	Val	Leu	Val
50						55					60				
Asp	Leu	Asn	Gly	Asn	His	Leu	Trp	Val	Asn	Cys	Glu	Gln	Gln	Tyr	Ser
65					70					75					80
Ser	Lys	Thr	Tyr	Gln	Ala	Pro	Phe	Cys	His	Ser	Thr	Gln	Cys	Ser	Arg
				85					90					95	
Ala	Asn	Thr	His	Gln	Cys	Leu	Ser	Cys	Pro	Ala	Ala	Ser	Arg	Pro	Gly
			100					105					110		
Cys	His	Lys	Asn	Thr	Cys	Gly	Leu	Met	Ser	Thr	Asn	Pro	Ile	Thr	Gln
		115					120					125			
Gln	Thr	Gly	Leu	Gly	Glu	Leu	Gly	Glu	Asp	Val	Leu	Ala	Ile	His	Ala
	130					135					140				
Thr	Gln	Gly	Ser	Thr	Gln	Gln	Leu	Gly	Pro	Leu	Val	Thr	Val	Pro	Gln
145					150					155					160
Phe	Leu	Phe	Ser	Cys	Ala	Pro	Ser	Phe	Leu	Val	Gln	Lys	Gly	Leu	Pro
				165					170					175	
Arg	Asn	Thr	Gln	Gly	Val	Ala	Gly	Leu	Gly	His	Ala	Pro	Ile	Ser	Leu
			180					185					190		
Pro	Asn	Gln	Leu	Ala	Ser	His	Phe	Gly	Leu	Gln	Arg	Gln	Phe	Thr	Thr
		195					200					205			
Cys	Leu	Ser	Arg	Tyr	Pro	Thr	Ser	Lys	Gly	Ala	Ile	Ile	Phe	Gly	Asp
	210					215					220				
Ala	Pro	Asn	Asn	Met	Arg	Gln	Phe	Gln	Asn	Gln	Asp	Ile	Phe	His	Asp
225					230					235					240
Leu	Ala	Phe	Thr	Pro	Leu	Thr	Ile	Thr	Leu	Gln	Gly	Glu	Tyr	Asn	Val
				245					250					255	
Arg	Val	Asn	Ser	Ile	Arg	Ile	Asn	Gln	His	Ser	Val	Phe	Pro	Leu	Asn
			260					265					270		
Lys	Ile	Ser	Ser	Thr	Ile	Val	Gly	Ser	Thr	Ser	Gly	Gly	Thr	Met	Ile
	275						280					285			
Ser	Thr	Ser	Thr	Pro	His	Met	Val	Leu	Gln	Gln	Ser	Val	Tyr	Gln	Ala
	290					295					300				
Phe	Thr	Gln	Val	Phe	Ala	Gln	Gln	Leu	Pro	Lys	Gln	Ala	Gln	Val	Lys
305					310					315					320
Ser	Val	Ala	Pro	Phe	Gly	Leu	Cys	Phe	Asn	Ser	Asn	Lys	Ile	Asn	Ala
				325					330					335	
Tyr	Pro	Ser	Val	Asp	Leu	Val	Met	Asp	Lys	Pro	Asn	Gly	Pro	Val	Trp
			340					345					350		
Arg	Ile	Ser	Gly	Glu	Asp	Leu	Met	Val	Gln	Ala	Gln	Pro	Gly	Val	Thr
		355					360					365			
Cys	Leu	Gly	Val	Met	Asn	Gly	Gly	Met	Gln	Pro	Arg	Ala	Glu	Ile	Thr
	370					375					380				
Leu	Gly	Ala	Arg	Gln	Leu	Glu	Glu	Asn	Leu	Val	Val	Phe	Asp	Leu	Ala
385					390					395					400
Arg	Ser	Arg	Val	Gly	Phe	Ser	Thr	Ser	Ser	Leu	His	Ser	His	Gly	Val
				405					410					415	
Lys	Cys	Ala	Asp	Leu	Phe	Asn	Phe	Ala	Asn	Ala					
			420					425							

&lt;210&gt; 97

&lt;211&gt; 814

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(720)

&lt;223&gt; 19kd zein

&lt;400&gt; 97

```

atg gca gcc aag att ttt gcc ctc ctt gcc ctc ctt gct ctt tca gca 48
Met Ala Ala Lys Ile Phe Ala Leu Leu Ala Leu Leu Ala Leu Ser Ala
  1             5             10             15

aac gtt gct acc gcg act att att cca caa tgc tca caa caa tac ctc 96
Asn Val Ala Thr Ala Thr Ile Ile Pro Gln Cys Ser Gln Gln Tyr Leu
             20             25             30

tct ccg gtg aca gcc gcg aga ttt gaa tac cca act ata caa tcc tac 144
Ser Pro Val Thr Ala Ala Arg Phe Glu Tyr Pro Thr Ile Gln Ser Tyr
             35             40             45

agg cta caa cag gcc atc gca gca agc atc tta cgg tcg tta gca ttg 192
Arg Leu Gln Gln Ala Ile Ala Ala Ser Ile Leu Arg Ser Leu Ala Leu
             50             55             60

act gtc caa caa cca tat gcc cta ttg caa caa cca tcc tta gtg aat 240
Thr Val Gln Gln Pro Tyr Ala Leu Leu Gln Gln Pro Ser Leu Val Asn
             65             70             75             80

cta tat ctc caa aga atc gta gca caa caa cta caa caa caa ttg ctt 288
Leu Tyr Leu Gln Arg Ile Val Ala Gln Gln Leu Gln Gln Gln Leu Leu
             85             90             95

cca aca atc aat gaa gta gtt gca gcg aac ctt gat gct tac ctc cag 336
Pro Thr Ile Asn Glu Val Val Ala Ala Asn Leu Asp Ala Tyr Leu Gln
             100            105            110

caa caa caa ttt ctt cca ttc aat caa cta gct ggg gtg aac cct gct 384
Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Gly Val Asn Pro Ala
             115            120            125

gct tac ttg cag gca caa cag cta cta cca ttc aac caa ctt gtc agg 432
Ala Tyr Leu Gln Ala Gln Gln Leu Leu Pro Phe Asn Gln Leu Val Arg
             130            135            140

agc cct gct gcc ttc tta ctg cag caa cag ttg ttg cca ttc cat cta 480
Ser Pro Ala Ala Phe Leu Leu Gln Gln Gln Leu Leu Pro Phe His Leu
             145            150            155            160

caa gtt gtg gca aac att gct gct ttc ttg caa caa caa caa ttg ctg 528
Gln Val Val Ala Asn Ile Ala Ala Phe Leu Gln Gln Gln Gln Leu Leu
             165            170            175

cca ttt tac cca cag gtt gtg gga aac att aac gcc ttc ttg caa cag 576
Pro Phe Tyr Pro Gln Val Val Gly Asn Ile Asn Ala Phe Leu Gln Gln
             180            185            190

caa cag ttg ctg cca ttc tac cca cag gat gtg gca aac aat gtc gcc 624
Gln Gln Leu Leu Pro Phe Tyr Pro Gln Asp Val Ala Asn Asn Val Ala
             195            200            205

ttc tta caa caa caa caa ttg ctg cca ttt agc caa ctt gct ttg acg 672
Phe Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala Leu Thr
             210            215            220

```

## 102

aat cct acc acc tta ttg cag cag ccc acc att ggt ggt gcc atc ttc 720  
 Asn Pro Thr Thr Leu Leu Gln Gln Pro Thr Ile Gly Gly Ala Ile Phe  
 225 230 235 240

tagatttttt atgctttata ctgtaataat aaagttctca tactgatatg tgcaacttct 780  
 cagtaataaa agattagaga tctatatattt atta 814

<210> 98

<211> 240

<212> PRT

<213> Zea mays

<400> 98

Met Ala Ala Lys Ile Phe Ala Leu Leu Ala Leu Leu Ala Leu Ser Ala  
 1 5 10 15

Asn Val Ala Thr Ala Thr Ile Ile Pro Gln Cys Ser Gln Gln Tyr Leu  
 20 25 30

Ser Pro Val Thr Ala Ala Arg Phe Glu Tyr Pro Thr Ile Gln Ser Tyr  
 35 40 45

Arg Leu Gln Gln Ala Ile Ala Ala Ser Ile Leu Arg Ser Leu Ala Leu  
 50 55 60

Thr Val Gln Gln Pro Tyr Ala Leu Leu Gln Gln Pro Ser Leu Val Asn  
 65 70 75 80

Leu Tyr Leu Gln Arg Ile Val Ala Gln Gln Leu Gln Gln Gln Leu Leu  
 85 90 95

Pro Thr Ile Asn Glu Val Val Ala Ala Asn Leu Asp Ala Tyr Leu Gln  
 100 105 110

Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Gly Val Asn Pro Ala  
 115 120 125

Ala Tyr Leu Gln Ala Gln Gln Leu Leu Pro Phe Asn Gln Leu Val Arg  
 130 135 140

Ser Pro Ala Ala Phe Leu Leu Gln Gln Gln Leu Leu Pro Phe His Leu  
 145 150 155 160

Gln Val Val Ala Asn Ile Ala Ala Phe Leu Gln Gln Gln Gln Leu Leu  
 165 170 175

Pro Phe Tyr Pro Gln Val Val Gly Asn Ile Asn Ala Phe Leu Gln Gln  
 180 185 190

Gln Gln Leu Leu Pro Phe Tyr Pro Gln Asp Val Ala Asn Asn Val Ala  
 195 200 205

Phe Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala Leu Thr  
 210 215 220

Asn Pro Thr Thr Leu Leu Gln Gln Pro Thr Ile Gly Gly Ala Ile Phe  
 225 230 235 240

<210> 99

<211> 705

<212> DNA

<213> Zea mays

<220>

<221> CDS

<222> (1)..(702)

<223> 19 kd zein B1

```

<400> 99
atg gca gcc aaa ata ttt tgc ctc ctt atg ctc ctt ggt ctt tct gca 48
Met Ala Ala Lys Ile Phe Cys Leu Leu Met Leu Leu Gly Leu Ser Ala
  1             5             10             15

agt gct gct acg gcg acc att ttc cca caa tgc tca caa gct cct ata 96
Ser Ala Ala Thr Ala Thr Ile Phe Pro Gln Cys Ser Gln Ala Pro Ile
             20             25             30

gct tcc ctt ctt ccc ccg tac ctc tca cca gcg gtg tct tcg gta tgt 144
Ala Ser Leu Leu Pro Pro Tyr Leu Ser Pro Ala Val Ser Ser Val Cys
             35             40             45

gaa aac cca att ctt caa ccc tat agg atc caa cag gca atc gca gct 192
Glu Asn Pro Ile Leu Gln Pro Tyr Arg Ile Gln Gln Ala Ile Ala Ala
             50             55             60

ggc atc tta cct tta tca ccc ttg ttc ctc caa caa tca tca gcc cta 240
Gly Ile Leu Pro Leu Ser Pro Leu Phe Leu Gln Gln Ser Ser Ala Leu
             65             70             75             80

tta cag cag tta cct ttg gtg cat tta ttg gca caa aac atc agg gca 288
Leu Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala
             85             90             95

caa caa cta caa caa ctt gtg cta gca aac ctt gct gcc tac tct cag 336
Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln
             100             105             110

caa caa cag ttt ctt cca ttc aac caa cta gct gca ttg aac tct gct 384
Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn Ser Ala
             115             120             125

tct tat ttg caa caa caa caa cta cca ttc agc cag cta tct gct gcc 432
Ser Tyr Leu Gln Gln Gln Gln Leu Pro Phe Ser Gln Leu Ser Ala Ala
             130             135             140

tac ccc cag caa ttt ctt cca ttc aac caa ctg aca gct ttg aac tct 480
Tyr Pro Gln Gln Phe Leu Pro Phe Asn Gln Leu Thr Ala Leu Asn Ser
             145             150             155             160

cct gct tat tta cag cag caa caa cta cta cca ttc agc cag cta gct 528
Pro Ala Tyr Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala
             165             170             175

ggt gtg agc cct gct acc ttc ttg aca caa cca caa ttg ttg ccg ttc 576
Gly Val Ser Pro Ala Thr Phe Leu Thr Gln Pro Gln Leu Leu Pro Phe
             180             185             190

tac cag cac gct gcg cct aac gct ggc acc ctc tta caa ctg caa caa 624
Tyr Gln His Ala Ala Pro Asn Ala Gly Thr Leu Leu Gln Leu Gln Gln
             195             200             205

ttg ctg cca ttc aac caa ctt gct ttg aca aac cca aca gca ttc tac 672
Leu Leu Pro Phe Asn Gln Leu Ala Leu Thr Asn Pro Thr Ala Phe Tyr
             210             215             220

caa caa ccc atc att ggt ggt gcc ctc ttt tag 705
Gln Gln Pro Ile Ile Gly Gly Ala Leu Phe
             225             230

<210> 100
<211> 234
<212> PRT
<213> Zea mays
<400> 100
Met Ala Ala Lys Ile Phe Cys Leu Leu Met Leu Leu Gly Leu Ser Ala
  1             5             10             15

```

## 104

Ser Ala Ala Thr Ala Thr Ile Phe Pro Gln Cys Ser Gln Ala Pro Ile  
 20 25 30  
 Ala Ser Leu Leu Pro Pro Tyr Leu Ser Pro Ala Val Ser Ser Val Cys  
 35 40 45  
 Glu Asn Pro Ile Leu Gln Pro Tyr Arg Ile Gln Gln Ala Ile Ala Ala  
 50 55 60  
 Gly Ile Leu Pro Leu Ser Pro Leu Phe Leu Gln Gln Ser Ser Ala Leu  
 65 70 75 80  
 Leu Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala  
 85 90 95  
 Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln  
 100 105 110  
 Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn Ser Ala  
 115 120 125  
 Ser Tyr Leu Gln Gln Gln Gln Leu Pro Phe Ser Gln Leu Ser Ala Ala  
 130 135 140  
 Tyr Pro Gln Gln Phe Leu Pro Phe Asn Gln Leu Thr Ala Leu Asn Ser  
 145 150 155 160  
 Pro Ala Tyr Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala  
 165 170 175  
 Gly Val Ser Pro Ala Thr Phe Leu Thr Gln Pro Gln Leu Leu Pro Phe  
 180 185 190  
 Tyr Gln His Ala Ala Pro Asn Ala Gly Thr Leu Leu Gln Leu Gln Gln  
 195 200 205  
 Leu Leu Pro Phe Asn Gln Leu Ala Leu Thr Asn Pro Thr Ala Phe Tyr  
 210 215 220  
 Gln Gln Pro Ile Ile Gly Gly Ala Leu Phe  
 225 230

&lt;210&gt; 101

&lt;211&gt; 804

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(801)

&lt;223&gt; 19 kd zein B2

&lt;400&gt; 101

atg gca gcc aaa ata ttt tgc ctc att atg ctc ctt ggt ctt tct gca 48  
 Met Ala Ala Lys Ile Phe Cys Leu Ile Met Leu Leu Gly Leu Ser Ala  
 1 5 10 15  
 agt gct gct acg gcg agc att ttc ccg caa tgc tca caa gct cct ata 96  
 Ser Ala Ala Thr Ala Ser Ile Phe Pro Gln Cys Ser Gln Ala Pro Ile  
 20 25 30  
 gct tcc ctt ctt ccc cca tac ctc tca cca gcg atg tct tca gta tgt 144  
 Ala Ser Leu Leu Pro Pro Tyr Leu Ser Pro Ala Met Ser Ser Val Cys  
 35 40 45  
 gaa aat cca att ctt cta ccc tac agg atc caa cag gca atc gca gca 192  
 Glu Asn Pro Ile Leu Leu Pro Tyr Arg Ile Gln Gln Ala Ile Ala Ala  
 50 55 60

## 105

```

ggc atc tta cct tta tca ccc ttg ttc ctc caa caa tca tca gcc cta 240
Gly Ile Leu Pro Leu Ser Pro Leu Phe Leu Gln Gln Ser Ser Ala Leu
65 70 75 80

tta cag cag tta cct ttg gtg cat tta ttg gca caa aac atc agg gca 288
Leu Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala
85 90 95

caa caa cta caa caa ctc gtg cta gca aac ctt gct gcc tac tct cag 336
Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln
100 105 110

caa cag cag tta cct ttg gtg cat ttg ttg gca caa aac atc agg gca 384
Gln Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala
115 120 125

caa caa cta caa caa ctc gtg cta gca aac ctt gct gcc tac tct cag 432
Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln
130 135 140

caa caa cag ttt ctg cca ttc aac caa cta gct gca ttg aac tct gct 480
Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn Ser Ala
145 150 155 160

gct tat ttg cag caa caa caa cta cta cca ttc agc cag cta gct gct 528
Ala Tyr Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala Ala
165 170 175

gcc tac ccc cgg caa ttt ctt cca ttc aac caa ctg gca gca ttg aac 576
Ala Tyr Pro Arg Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn
180 185 190

tct cat gct tat gta caa caa caa caa cta cta cca ttc agc cag cta 624
Ser His Ala Tyr Val Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu
195 200 205

gct gct gtg agc cct gct gcc ttc ttg aca cag caa cat ttg ttg ccg 672
Ala Ala Val Ser Pro Ala Ala Phe Leu Thr Gln Gln His Leu Leu Pro
210 215 220

ttc tac ctg cac act gcg cct aac gtt ggc acc ctc tta caa ctg caa 720
Phe Tyr Leu His Thr Ala Pro Asn Val Gly Thr Leu Leu Gln Leu Gln
225 230 235 240

caa ttg ctg cca ttc gac caa ctt gct ttg aca aac cca gca gtg ttc 768
Gln Leu Leu Pro Phe Asp Gln Leu Ala Leu Thr Asn Pro Ala Val Phe
245 250 255

tac caa caa ccc atc att ggt ggt gcc ctc ttt tag 804
Tyr Gln Gln Pro Ile Ile Gly Gly Ala Leu Phe
260 265

```

&lt;210&gt; 102

&lt;211&gt; 267

&lt;212&gt; PRT

&lt;213&gt; Zea mays

&lt;400&gt; 102

```

Met Ala Ala Lys Ile Phe Cys Leu Ile Met Leu Leu Gly Leu Ser Ala
1 5 10 15

Ser Ala Ala Thr Ala Ser Ile Phe Pro Gln Cys Ser Gln Ala Pro Ile
20 25 30

Ala Ser Leu Leu Pro Pro Tyr Leu Ser Pro Ala Met Ser Ser Val Cys
35 40 45

```



## 106

Glu Asn Pro Ile Leu Leu Pro Tyr Arg Ile Gln Gln Ala Ile Ala Ala  
     50                    55                    60  
 Gly Ile Leu Pro Leu Ser Pro Leu Phe Leu Gln Gln Ser Ser Ala Leu  
     65                    70                    75                    80  
 Leu Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala  
                     85                    90                    95  
 Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln  
                     100                    105                    110  
 Gln Gln Gln Leu Pro Leu Val His Leu Leu Ala Gln Asn Ile Arg Ala  
                     115                    120                    125  
 Gln Gln Leu Gln Gln Leu Val Leu Ala Asn Leu Ala Ala Tyr Ser Gln  
     130                    135                    140  
 Gln Gln Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn Ser Ala  
     145                    150                    155                    160  
 Ala Tyr Leu Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu Ala Ala  
                     165                    170                    175  
 Ala Tyr Pro Arg Gln Phe Leu Pro Phe Asn Gln Leu Ala Ala Leu Asn  
                     180                    185                    190  
 Ser His Ala Tyr Val Gln Gln Gln Gln Leu Leu Pro Phe Ser Gln Leu  
                     195                    200                    205  
 Ala Ala Val Ser Pro Ala Ala Phe Leu Thr Gln Gln His Leu Leu Pro  
     210                    215                    220  
 Phe Tyr Leu His Thr Ala Pro Asn Val Gly Thr Leu Leu Gln Leu Gln  
     225                    230                    235                    240  
 Gln Leu Leu Pro Phe Asp Gln Leu Ala Leu Thr Asn Pro Ala Val Phe  
                     245                    250                    255  
 Tyr Gln Gln Pro Ile Ile Gly Gly Ala Leu Phe  
                     260                    265

&lt;210&gt; 103

&lt;211&gt; 801

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(798)

&lt;223&gt; 22kd alpha-zein

&lt;400&gt; 103

atg gct acc aag ata tta gcc ctc ctt gcg ctc ctt tcc ctt tca gtg 48  
 Met Ala Thr Lys Ile Leu Ala Leu Leu Ala Leu Leu Ser Leu Ser Val  
     1                    5                    10                    15  
 agc gca aca act gca ttc att att cca caa tgc tca ctt gct cct aat 96  
 Ser Ala Thr Thr Ala Phe Ile Ile Pro Gln Cys Ser Leu Ala Pro Asn  
                     20                    25                    30  
 gcc att att cca cag ttc ctc cca tca gtt aca tca atg ggc atc gaa 144  
 Ala Ile Ile Pro Gln Phe Leu Pro Ser Val Thr Ser Met Gly Ile Glu  
                     35                    40                    45  
 cac cct att gtg caa gcc tat agg cta caa caa gcg ctt gcg gcg agc 192  
 His Pro Ile Val Gln Ala Tyr Arg Leu Gln Gln Ala Leu Ala Ala Ser  
     50                    55                    60

## 107

```

gtc tta caa caa ccg ttt gcc caa tta caa caa caa tcc ttg gca cat 240
Val Leu Gln Gln Pro Phe Ala Gln Leu Gln Gln Gln Ser Leu Ala His
65 70 75 80

cta acc ata caa acc atc gca aca caa cta gag caa cag ttt gtg ccc 288
Leu Thr Ile Gln Thr Ile Ala Thr Gln Leu Glu Gln Gln Phe Val Pro
85 90 95

gca ttg agc caa cta gcc gcg gtg aac cct gtc tcc tac ttg caa cag 336
Ala Leu Ser Gln Leu Ala Ala Val Asn Pro Val Ser Tyr Leu Gln Gln
100 105 110

caa atg ctt gca tcc aac cca ctt gct ctg gcg aac aca gcc gca tac 384
Gln Met Leu Ala Ser Asn Pro Leu Ala Leu Ala Asn Thr Ala Ala Tyr
115 120 125

cag caa caa cta cag ttg caa cag ttt cta cca gct ctt agt caa cta 432
Gln Gln Gln Leu Gln Leu Gln Gln Phe Leu Pro Ala Leu Ser Gln Leu
130 135 140

gcc agg gtg aac cct gcc aca tac ctg caa cag caa caa ctg ctt tca 480
Ala Arg Val Asn Pro Ala Thr Tyr Leu Gln Gln Gln Leu Leu Ser
145 150 155 160

tct agc cca ctc gct gtg ggc aat gcg gct aca tac ctg caa cag cag 528
Ser Ser Pro Leu Ala Val Gly Asn Ala Ala Thr Tyr Leu Gln Gln Gln
165 170 175

ctg cta caa cag atc gta ccg gct ctt agt cag cta gtt gtg gcg aac 576
Leu Leu Gln Gln Ile Val Pro Ala Leu Ser Gln Leu Val Val Ala Asn
180 185 190

cct act gcc tac tta caa cag ctt ctt cca ttc aac caa cta gat gtg 624
Pro Thr Ala Tyr Leu Gln Gln Leu Leu Pro Phe Asn Gln Leu Asp Val
195 200 205

gca aac tct gct gcg tac cta caa cag cgg cag caa cta ctt aat cca 672
Ala Asn Ser Ala Ala Tyr Leu Gln Gln Arg Gln Gln Leu Leu Asn Pro
210 215 220

ctt gca gcg gct aac cca ttg gtc gcc gcc ttc ctg caa cag caa caa 720
Leu Ala Ala Ala Asn Pro Leu Val Ala Ala Phe Leu Gln Gln Gln Gln
225 230 235 240

ttt ctg cca tac aac caa atc tct ttg atg aac ctt gcc ttg tca agg 768
Phe Leu Pro Tyr Asn Gln Ile Ser Leu Met Asn Leu Ala Leu Ser Arg
245 250 255

cag caa ccg atc gtt gga ggt gcc atc ttt tag 801
Gln Gln Pro Ile Val Gly Gly Ala Ile Phe
260 265

```

&lt;210&gt; 104

&lt;211&gt; 266

&lt;212&gt; PRT

&lt;213&gt; Zea mays

&lt;400&gt; 104

```

Met Ala Thr Lys Ile Leu Ala Leu Leu Ala Leu Leu Ser Leu Ser Val
1 5 10 15

```

```

Ser Ala Thr Thr Ala Phe Ile Ile Pro Gln Cys Ser Leu Ala Pro Asn
20 25 30

```

```

Ala Ile Ile Pro Gln Phe Leu Pro Ser Val Thr Ser Met Gly Ile Glu
35 40 45

```

## 108

His Pro Ile Val Gln Ala Tyr Arg Leu Gln Gln Ala Leu Ala Ala Ser  
 50 55 60  
 Val Leu Gln Gln Pro Phe Ala Gln Leu Gln Gln Gln Ser Leu Ala His  
 65 70 75 80  
 Leu Thr Ile Gln Thr Ile Ala Thr Gln Leu Glu Gln Gln Phe Val Pro  
 85 90 95  
 Ala Leu Ser Gln Leu Ala Ala Val Asn Pro Val Ser Tyr Leu Gln Gln  
 100 105 110  
 Gln Met Leu Ala Ser Asn Pro Leu Ala Leu Ala Asn Thr Ala Ala Tyr  
 115 120 125  
 Gln Gln Gln Leu Gln Leu Gln Gln Phe Leu Pro Ala Leu Ser Gln Leu  
 130 135 140  
 Ala Arg Val Asn Pro Ala Thr Tyr Leu Gln Gln Gln Gln Leu Leu Ser  
 145 150 155 160  
 Ser Ser Pro Leu Ala Val Gly Asn Ala Ala Thr Tyr Leu Gln Gln Gln  
 165 170 175  
 Leu Leu Gln Gln Ile Val Pro Ala Leu Ser Gln Leu Val Val Ala Asn  
 180 185 190  
 Pro Thr Ala Tyr Leu Gln Gln Leu Leu Pro Phe Asn Gln Leu Asp Val  
 195 200 205  
 Ala Asn Ser Ala Ala Tyr Leu Gln Gln Arg Gln Gln Leu Leu Asn Pro  
 210 215 220  
 Leu Ala Ala Ala Asn Pro Leu Val Ala Ala Phe Leu Gln Gln Gln Gln  
 225 230 235 240  
 Phe Leu Pro Tyr Asn Gln Ile Ser Leu Met Asn Leu Ala Leu Ser Arg  
 245 250 255  
 Gln Gln Pro Ile Val Gly Gly Ala Ile Phe  
 260 265

&lt;210&gt; 105

&lt;211&gt; 471

&lt;212&gt; DNA

&lt;213&gt; Oryza sativa

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(468)

&lt;223&gt; prolamin

&lt;400&gt; 105

atg aag atc att ttc gta ttt gct ctc ctt gct att gtt gca tgc aac 48  
 Met Lys Ile Ile Phe Val Phe Ala Leu Leu Ala Ile Val Ala Cys Asn  
 1 5 10 15

gct tct gca cgg ttt gat gct ctt agt caa agt tat aga caa tat caa 96  
 Ala Ser Ala Arg Phe Asp Ala Leu Ser Gln Ser Tyr Arg Gln Tyr Gln  
 20 25 30

cta caa tcg cat ctc ctg cta cag caa caa gtg ctc agc cca tgc agt 144  
 Leu Gln Ser His Leu Leu Leu Gln Gln Gln Val Leu Ser Pro Cys Ser  
 35 40 45

gag ttc gta agg caa cag cat agc ata gtg gca acc ccc ttc tgg caa 192  
 Glu Phe Val Arg Gln Gln His Ser Ile Val Ala Thr Pro Phe Trp Gln  
 50 55 60

## 109

```

cca gct acg ttt caa ttg ata aac aac caa gtc atg cag caa cag tgt      240
Pro Ala Thr Phe Gln Leu Ile Asn Asn Gln Val Met Gln Gln Gln Cys
  65                      70                      75                      80

tgc caa cag ctc agg ctg gta gcg caa caa tct cac tac cag gcc att      288
Cys Gln Gln Leu Arg Leu Val Ala Gln Gln Ser His Tyr Gln Ala Ile
                      85                      90                      95

agt agc gtt cag gcg att gtg cag caa cta cag ctg cag cag gtc ggt      336
Ser Ser Val Gln Ala Ile Val Gln Gln Leu Gln Leu Gln Gln Val Gly
                      100                      105                      110

gtt gtc tac ttt gat cag act caa gct caa gct caa gct ttg ctg gcc      384
Val Val Tyr Phe Asp Gln Thr Gln Ala Gln Ala Gln Ala Leu Leu Ala
                      115                      120                      125

tta aac ttg cca tcc ata tgt ggt atc tat cct aac tac tac att gct      432
Leu Asn Leu Pro Ser Ile Cys Gly Ile Tyr Pro Asn Tyr Tyr Ile Ala
                      130                      135                      140

ccg agg agc att ccc acc gtt ggt ggt gtc tgg tac tga                  471
Pro Arg Ser Ile Pro Thr Val Gly Gly Val Trp Tyr
145                      150                      155

```

&lt;210&gt; 106

&lt;211&gt; 156

&lt;212&gt; PRT

&lt;213&gt; Oryza sativa

&lt;400&gt; 106

```

Met Lys Ile Ile Phe Val Phe Ala Leu Leu Ala Ile Val Ala Cys Asn
  1                      5                      10                      15

Ala Ser Ala Arg Phe Asp Ala Leu Ser Gln Ser Tyr Arg Gln Tyr Gln
                      20                      25                      30

Leu Gln Ser His Leu Leu Leu Gln Gln Gln Val Leu Ser Pro Cys Ser
                      35                      40                      45

Glu Phe Val Arg Gln Gln His Ser Ile Val Ala Thr Pro Phe Trp Gln
                      50                      55                      60

Pro Ala Thr Phe Gln Leu Ile Asn Asn Gln Val Met Gln Gln Gln Cys
  65                      70                      75                      80

Cys Gln Gln Leu Arg Leu Val Ala Gln Gln Ser His Tyr Gln Ala Ile
                      85                      90                      95

Ser Ser Val Gln Ala Ile Val Gln Gln Leu Gln Leu Gln Gln Val Gly
                      100                      105                      110

Val Val Tyr Phe Asp Gln Thr Gln Ala Gln Ala Gln Ala Leu Leu Ala
                      115                      120                      125

Leu Asn Leu Pro Ser Ile Cys Gly Ile Tyr Pro Asn Tyr Tyr Ile Ala
                      130                      135                      140

Pro Arg Ser Ile Pro Thr Val Gly Gly Val Trp Tyr
145                      150                      155

```

&lt;210&gt; 107

&lt;211&gt; 645

&lt;212&gt; DNA

&lt;213&gt; Avena sativa

&lt;220&gt;

&lt;221&gt; CDS

## 110

&lt;222&gt; (1)..(642)

&lt;223&gt; avenin

&lt;400&gt; 107

atg	aag	atc	ttc	ttc	ttc	tta	gct	ctc	ctt	gct	ctg	gta	gtg	agc	gcc	48
Met	Lys	Ile	Phe	Phe	Phe	Leu	Ala	Leu	Leu	Ala	Leu	Val	Val	Ser	Ala	
1				5					10					15		

acc	ttt	gca	caa	tat	gca	gaa	tct	gac	ggt	agt	tat	gag	gaa	gtg	gag	96
Thr	Phe	Ala	Gln	Tyr	Ala	Glu	Ser	Asp	Gly	Ser	Tyr	Glu	Glu	Val	Glu	
			20					25					30			

ggt	tct	cat	gat	cga	tgc	caa	caa	cat	cag	atg	aag	ctg	gac	tct	tgc	144
Gly	Ser	His	Asp	Arg	Cys	Gln	Gln	His	Gln	Met	Lys	Leu	Asp	Ser	Cys	
		35					40					45				

aga	gag	tac	gtg	gcg	gag	cgg	tgc	aca	acg	atg	aga	gat	ttt	ccg	atc	192
Arg	Glu	Tyr	Val	Ala	Glu	Arg	Cys	Thr	Thr	Met	Arg	Asp	Phe	Pro	Ile	
	50					55					60					

acc	tgg	cca	tgg	aaa	tgg	tgg	aag	ggt	ggt	tgc	gag	gag	ctc	cgc	aat	240
Thr	Trp	Pro	Trp	Lys	Trp	Trp	Lys	Gly	Gly	Cys	Glu	Glu	Leu	Arg	Asn	
65				70					75						80	

gag	tgc	tgc	caa	ctg	ttg	ggc	cag	atg	cca	tgc	gag	tgt	cgc	tgt	gat	288
Glu	Cys	Cys	Gln	Leu	Leu	Gly	Gln	Met	Pro	Ser	Glu	Cys	Arg	Cys	Asp	
			85					90					95			

gcg	att	tgg	aga	tca	atc	cag	cgc	gag	ctt	ggt	ggc	ttc	ttt	gga	act	336
Ala	Ile	Trp	Arg	Ser	Ile	Gln	Arg	Glu	Leu	Gly	Gly	Phe	Phe	Gly	Thr	
			100					105					110			

caa	caa	ggt	ctg	ata	ggg	aaa	agg	ttg	aag	ata	gcc	aag	agt	ttg	ccc	384
Gln	Gln	Gly	Leu	Ile	Gly	Lys	Arg	Leu	Lys	Ile	Ala	Lys	Ser	Leu	Pro	
		115					120					125				

acg	cag	tca	aca	tgg	gcc	ctg	agt	gca	ata	tcc	cca	aac	tcc	atg	gtt	432
Thr	Gln	Ser	Thr	Trp	Ala	Leu	Ser	Ala	Ile	Ser	Pro	Asn	Ser	Met	Val	
	130					135					140					

agc	cac	att	gct	gga	aag	agc	tcc	att	ctt	cgt	gcc	ttg	ccc	gtg	gat	480
Ser	His	Ile	Ala	Gly	Lys	Ser	Ser	Ile	Leu	Arg	Ala	Leu	Pro	Val	Asp	
145					150					155				160		

gtc	ctc	gcc	aat	gca	tac	cgc	att	tcc	agg	caa	gaa	gcc	cga	aac	ctc	528
Val	Leu	Ala	Asn	Ala	Tyr	Arg	Ile	Ser	Arg	Gln	Glu	Ala	Arg	Asn	Leu	
			165					170					175			

aaa	aac	aac	agg	gga	caa	gag	tct	ggt	gta	ttc	act	cca	aaa	ttt	acc	576
Lys	Asn	Asn	Arg	Gly	Gln	Glu	Ser	Gly	Val	Phe	Thr	Pro	Lys	Phe	Thr	
			180					185					190			

caa	acg	agc	ttc	caa	cct	tat	cca	gag	ggc	gag	gat	gag	tca	tct	ttg	624
Gln	Thr	Ser	Phe	Gln	Pro	Tyr	Pro	Glu	Gly	Glu	Asp	Glu	Ser	Ser	Leu	
		195					200					205				

att	aat	aag	gca	tca	gag	taa										645
Ile	Asn	Lys	Ala	Ser	Glu											

&lt;210&gt; 108

&lt;211&gt; 214

&lt;212&gt; PRT

&lt;213&gt; Avena sativa

&lt;400&gt; 108

Met	Lys	Ile	Phe	Phe	Phe	Leu	Ala	Leu	Leu	Ala	Leu	Val	Val	Ser	Ala
1				5					10					15	

## 111

Thr Phe Ala Gln Tyr Ala Glu Ser Asp Gly Ser Tyr Glu Glu Val Glu  
                     20                    25                    30  
 Gly Ser His Asp Arg Cys Gln Gln His Gln Met Lys Leu Asp Ser Cys  
                     35                    40                    45  
 Arg Glu Tyr Val Ala Glu Arg Cys Thr Thr Met Arg Asp Phe Pro Ile  
                     50                    55                    60  
 Thr Trp Pro Trp Lys Trp Trp Lys Gly Gly Cys Glu Glu Leu Arg Asn  
                     65                    70                    75                    80  
 Glu Cys Cys Gln Leu Leu Gly Gln Met Pro Ser Glu Cys Arg Cys Asp  
                     85                    90                    95  
 Ala Ile Trp Arg Ser Ile Gln Arg Glu Leu Gly Gly Phe Phe Gly Thr  
                     100                    105                    110  
 Gln Gln Gly Leu Ile Gly Lys Arg Leu Lys Ile Ala Lys Ser Leu Pro  
                     115                    120                    125  
 Thr Gln Ser Thr Trp Ala Leu Ser Ala Ile Ser Pro Asn Ser Met Val  
                     130                    135                    140  
 Ser His Ile Ala Gly Lys Ser Ser Ile Leu Arg Ala Leu Pro Val Asp  
                     145                    150                    155                    160  
 Val Leu Ala Asn Ala Tyr Arg Ile Ser Arg Gln Glu Ala Arg Asn Leu  
                     165                    170                    175  
 Lys Asn Asn Arg Gly Gln Glu Ser Gly Val Phe Thr Pro Lys Phe Thr  
                     180                    185                    190  
 Gln Thr Ser Phe Gln Pro Tyr Pro Glu Gly Glu Asp Glu Ser Ser Leu  
                     195                    200                    205  
 Ile Asn Lys Ala Ser Glu  
                     210

&lt;210&gt; 109

&lt;211&gt; 1044

&lt;212&gt; DNA

&lt;213&gt; Hordeum vulgare

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1041)

&lt;223&gt; c-hordein

&lt;220&gt;

&lt;221&gt; misc\_feature

&lt;222&gt; (481)..(482)

&lt;223&gt; /transl\_except=(pos:481..483,aa:OTHER)

&lt;400&gt; 109

atg aag acg ttc ctc acc ttt gtc ctc ctt gcc atg gcg atg agc atc 48  
 Met Lys Thr Phe Leu Thr Phe Val Leu Leu Ala Met Ala Met Ser Ile  
                     1                    5                    10                    15

gtc act acc gct agg cag cta aac cct agc cac caa gag ttg caa tca 96  
 Val Thr Thr Ala Arg Gln Leu Asn Pro Ser His Gln Glu Leu Gln Ser  
                     20                    25                    30

cca caa caa cca ttt ctg aaa caa caa tca tat ctg caa caa cca tat 144  
 Pro Gln Gln Pro Phe Leu Lys Gln Gln Ser Tyr Leu Gln Gln Pro Tyr  
                     35                    40                    45

## 112

cca	caa	caa	cca	tat	cta	ccg	cag	caa	cca	ttc	ccc	aca	ccc	caa	caa	192
Pro	Gln	Gln	Pro	Tyr	Leu	Pro	Gln	Gln	Pro	Phe	Pro	Thr	Pro	Gln	Gln	
50						55					60					
ttt	ttc	ccc	tat	cta	cca	cag	caa	aca	ttt	ccc	cca	tcc	caa	caa	cca	240
Phe	Phe	Pro	Tyr	Leu	Pro	Gln	Gln	Thr	Phe	Pro	Pro	Ser	Gln	Gln	Pro	
65					70					75					80	
aac	ccc	cta	caa	cca	caa	caa	cca	ttc	ccc	ctg	caa	ccc	caa	cca	cca	288
Asn	Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	Gln	Pro	Pro	
				85						90				95		
caa	caa	cct	ttt	cct	cag	ccc	caa	caa	cca	aat	ccc	cag	caa	cca	caa	336
Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Gln	Pro	Asn	Pro	Gln	Gln	Pro	Gln	
			100					105					110			
caa	cct	ttc	ccc	cgg	caa	cca	caa	caa	ata	gta	ccc	cag	caa	cca	caa	384
Gln	Pro	Phe	Pro	Arg	Gln	Pro	Gln	Gln	Ile	Val	Pro	Gln	Gln	Pro	Gln	
		115					120					125				
caa	cca	ttc	cct	cag	caa	cca	caa	caa	cct	ttt	cct	cag	ccc	caa	caa	432
Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Gln	
		130				135					140					
cca	ttc	tct	tg	caa	cca	caa	caa	cca	ttt	ctc	cag	ccc	cta	caa	cta	480
Pro	Phe	Ser	Trp	Gln	Pro	Gln	Gln	Pro	Phe	Leu	Gln	Pro	Leu	Gln	Leu	
145					150					155					160	
tag	ccc	ctg	caa	gca	caa	caa	cca	ttc	ccc	ttg	caa	cct	caa	cta	cca	528
	Pro	Leu	Gln	Ala	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	Gln	Leu	Pro	
				165						170				175		
ttt	ccg	caa	ccc	caa	caa	cca	att	gga	cag	caa	cca	aaa	caa	cca	ctc	576
Phe	Pro	Gln	Pro	Gln	Gln	Pro	Ile	Gly	Gln	Gln	Pro	Lys	Gln	Pro	Leu	
			180					185					190			
ctg	cag	caa	cca	caa	caa	aca	att	ccc	cag	caa	cca	caa	caa	cca	ttc	624
Leu	Gln	Gln	Pro	Gln	Gln	Thr	Ile	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Phe	
			195				200					205				
ccc	ctg	cag	ccg	caa	caa	cca	ttc	ccc	caa	caa	cca	caa	caa	cca	ctt	672
Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	
			210			215					220					
ccc	caa	caa	ccc	caa	caa	ata	att	tcc	cag	caa	ccc	caa	caa	cca	ttc	720
Pro	Gln	Gln	Pro	Gln	Gln	Ile	Ile	Ser	Gln	Gln	Pro	Gln	Gln	Pro	Phe	
225					230					235					240	
cct	cta	caa	cct	caa	caa	cca	ttc	ccc	caa	ccc	caa	cca	ttc	ccc	cag	768
Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	Gln	Pro	Gln	Pro	Phe	Pro	Gln	
				245					250					255		
gag	caa	ccc	caa	caa	gca	ttc	ccc	cta	caa	ccg	caa	caa	cca	ttc	ccc	816
Glu	Gln	Pro	Gln	Gln	Ala	Phe	Pro	Leu	Gln	Pro	Gln	Gln	Pro	Phe	Pro	
				260				265						270		
gag	gaa	tca	gaa	caa	ata	att	acc	caa	caa	cca	ttc	cct	cta	caa	cca	864
Glu	Glu	Ser	Glu	Gln	Ile	Ile	Thr	Gln	Gln	Pro	Phe	Pro	Leu	Gln	Pro	
			275				280						285			
caa	caa	ctg	ttc	ccc	cag	caa	cca	caa	caa	cca	ctt	ccc	cag	ccc	caa	912
Gln	Gln	Leu	Phe	Pro	Gln	Gln	Pro	Gln	Gln	Pro	Leu	Pro	Gln	Pro	Gln	
			290			295					300					
caa	cca	ttc	cgc	caa	cta	cca	aaa	tat	ata	att	ccc	cag	caa	cct	caa	960
Gln	Pro	Phe	Arg	Gln	Leu	Pro	Lys	Tyr	Ile	Ile	Pro	Gln	Gln	Pro	Gln	
305					310					315					320	

## 113

caa cca ttc ctt ctg caa cca cac caa cct cag caa cct tat gca caa 1008  
 Gln Pro Phe Leu Leu Gln Pro His Gln Pro Gln Gln Pro Tyr Ala Gln  
 325 330 335

caa gac atc tgg agt gat ata gcc ctc ttg ggc taa 1044  
 Gln Asp Ile Trp Ser Asp Ile Ala Leu Leu Gly  
 340 345

<210> 110

<211> 160

<212> PRT

<213> Hordeum vulgare

<400> 110

Met Lys Thr Phe Leu Thr Phe Val Leu Leu Ala Met Ala Met Ser Ile  
 1 5 10 15  
 Val Thr Thr Ala Arg Gln Leu Asn Pro Ser His Gln Glu Leu Gln Ser  
 20 25 30  
 Pro Gln Gln Pro Phe Leu Lys Gln Ser Tyr Leu Gln Gln Pro Tyr  
 35 40 45  
 Pro Gln Gln Pro Tyr Leu Pro Gln Gln Pro Phe Pro Thr Pro Gln Gln  
 50 55 60  
 Phe Phe Pro Tyr Leu Pro Gln Gln Thr Phe Pro Pro Ser Gln Gln Pro  
 65 70 75 80  
 Asn Pro Leu Gln Pro Gln Gln Pro Phe Pro Leu Gln Pro Gln Pro Pro  
 85 90 95  
 Gln Gln Pro Phe Pro Gln Pro Gln Gln Pro Asn Pro Gln Gln Pro Gln  
 100 105 110  
 Gln Pro Phe Pro Arg Gln Pro Gln Gln Ile Val Pro Gln Gln Pro Gln  
 115 120 125  
 Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Phe Pro Gln Pro Gln Gln  
 130 135 140  
 Pro Phe Ser Trp Gln Pro Gln Gln Pro Phe Leu Gln Pro Leu Gln Leu  
 145 150 155 160

<210> 111

<211> 186

<212> PRT

<213> Hordeum vulgare

<400> 111

Pro Leu Gln Ala Gln Gln Pro Phe Pro Leu Gln Pro Gln Leu Pro Phe  
 1 5 10 15  
 Pro Gln Pro Gln Gln Pro Ile Gly Gln Gln Pro Lys Gln Pro Leu Leu  
 20 25 30  
 Gln Gln Pro Gln Gln Thr Ile Pro Gln Gln Pro Gln Gln Pro Phe Pro  
 35 40 45  
 Leu Gln Pro Gln Gln Pro Phe Pro Gln Gln Pro Gln Gln Pro Leu Pro  
 50 55 60  
 Gln Gln Pro Gln Gln Ile Ile Ser Gln Gln Pro Gln Gln Pro Phe Pro  
 65 70 75 80  
 Leu Gln Pro Gln Gln Pro Phe Pro Gln Pro Gln Pro Phe Pro Gln Glu  
 85 90 95  
 Gln Pro Gln Gln Ala Phe Pro Leu Gln Pro Gln Gln Pro Phe Pro Glu  
 100 105 110  
 Glu Ser Glu Gln Ile Ile Thr Gln Gln Pro Phe Pro Leu Gln Pro Gln  
 115 120 125  
 Gln Leu Phe Pro Gln Gln Pro Gln Gln Pro Leu Pro Gln Pro Gln Gln  
 130 135 140  
 Pro Phe Arg Gln Leu Pro Lys Tyr Ile Ile Pro Gln Gln Pro Gln Gln  
 145 150 155 160



## 114

Pro Phe Leu Leu Gln Pro His Gln Pro Gln Gln Pro Tyr Ala Gln Gln  
 165 170 175  
 Asp Ile Trp Ser Asp Ile Ala Leu Leu Gly  
 180 185

<210> 112  
 <211> 924  
 <212> DNA  
 <213> Triticum aestivum  
 <220>  
 <221> CDS  
 <222> (1)..(921)  
 <223> glutenin-1D1

<400> 112  
 atg aag acc ttc ctc gtc ttt gcc ctc ctc gcc gtt gcg gcg aca agt 48  
 Met Lys Thr Phe Leu Val Phe Ala Leu Leu Ala Val Ala Ala Thr Ser  
 1 5 10 15  
 gca att gcg cag atg gag act aga tgc atc cct ggt ttg gag aga cca 96  
 Ala Ile Ala Gln Met Glu Thr Arg Cys Ile Pro Gly Leu Glu Arg Pro  
 20 25 30  
 tgg cag cag caa cca tta cca cca caa cag aca ttt cca caa caa cca 144  
 Trp Gln Gln Gln Pro Leu Pro Pro Gln Gln Thr Phe Pro Gln Gln Pro  
 35 40 45  
 cta ttt tca caa caa caa caa caa caa cta ttt cct caa caa cca tca 192  
 Leu Phe Ser Gln Gln Gln Gln Gln Gln Leu Phe Pro Gln Gln Pro Ser  
 50 55 60  
 ttt tcg cag caa caa cca cca ttt tgg cag caa caa cca cca ttt tct 240  
 Phe Ser Gln Gln Gln Pro Pro Phe Trp Gln Gln Gln Pro Pro Phe Ser  
 65 70 75 80  
 cag caa caa cca att cta cca cag caa cca cca ttt tcg cag caa caa 288  
 Gln Gln Gln Pro Ile Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln  
 85 90 95  
 caa cta gtt cta ccg caa caa cca cca ttt tca cag caa caa caa cca 336  
 Gln Leu Val Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln Gln Pro  
 100 105 110  
 gtt tta cct cca caa caa tca cct ttt cca caa caa caa caa caa cac 384  
 Val Leu Pro Pro Gln Gln Ser Pro Phe Pro Gln Gln Gln Gln Gln His  
 115 120 125  
 caa cag ctg gtg caa caa caa atc cct gtt gtt cag cca tcc att ttg 432  
 Gln Gln Leu Val Gln Gln Gln Ile Pro Val Val Gln Pro Ser Ile Leu  
 130 135 140  
 cag cag cta aac cca tgc aag gta ttc ctc cag cag cag tgc agc cct 480  
 Gln Gln Leu Asn Pro Cys Lys Val Phe Leu Gln Gln Gln Cys Ser Pro  
 145 150 155 160  
 gtg gca atg cca caa cgt ctt gct agg tcg caa atg ttg cag cag agc 528  
 Val Ala Met Pro Gln Arg Leu Ala Arg Ser Gln Met Leu Gln Gln Ser  
 165 170 175  
 agt tgc cat gtg atg caa caa caa tgt tgc cag cag ttg ccg caa atc 576  
 Ser Cys His Val Met Gln Gln Gln Cys Cys Gln Gln Leu Pro Gln Ile  
 180 185 190  
 ccc cag caa tcc cgc tat gag gca atc cgt gct atc atc tac tcc atc 624  
 Pro Gln Gln Ser Arg Tyr Glu Ala Ile Arg Ala Ile Ile Tyr Ser Ile  
 195 200 205

## 115

atc ctg caa gaa caa caa cag gtt cag ggt tcc atc caa tct cag cag 672  
 Ile Leu Gln Glu Gln Gln Gln Val Gln Gly Ser Ile Gln Ser Gln Gln  
 210 215 220  
 cag caa ccc caa cag ttg ggc caa tgt gtt tcc caa ccc caa cag cag 720  
 Gln Gln Pro Gln Gln Leu Gly Gln Cys Val Ser Gln Pro Gln Gln Gln  
 225 230 235 240  
 tcg cag cag caa ctc ggg caa caa cct caa caa caa caa ttg gca cag 768  
 Ser Gln Gln Gln Leu Gly Gln Gln Pro Gln Gln Gln Gln Leu Ala Gln  
 245 250 255  
 ggt acc ttt ttg cag cca cac cag ata gct cag ctt gag gtg atg act 816  
 Gly Thr Phe Leu Gln Pro His Gln Ile Ala Gln Leu Glu Val Met Thr  
 260 265 270  
 tcc att gcg ctc cgt atc ctg cca acg atg tgc agt gtt aat gtg ccg 864  
 Ser Ile Ala Leu Arg Ile Leu Pro Thr Met Cys Ser Val Asn Val Pro  
 275 280 285  
 ttg tac aga acc acc act agt gtg cca ttc ggc gtt ggc acc gga gtt 912  
 Leu Tyr Arg Thr Thr Thr Ser Val Pro Phe Gly Val Gly Thr Gly Val  
 290 295 300  
 ggt gcc tac tga 924  
 Gly Ala Tyr  
 305

&lt;210&gt; 113

&lt;211&gt; 307

&lt;212&gt; PRT

&lt;213&gt; Triticum aestivum

&lt;400&gt; 113

Met Lys Thr Phe Leu Val Phe Ala Leu Leu Ala Val Ala Ala Thr Ser  
 1 5 10 15  
 Ala Ile Ala Gln Met Glu Thr Arg Cys Ile Pro Gly Leu Glu Arg Pro  
 20 25 30  
 Trp Gln Gln Gln Pro Leu Pro Pro Gln Gln Thr Phe Pro Gln Gln Pro  
 35 40 45  
 Leu Phe Ser Gln Gln Gln Gln Gln Gln Leu Phe Pro Gln Gln Pro Ser  
 50 55 60  
 Phe Ser Gln Gln Gln Pro Pro Phe Trp Gln Gln Gln Pro Pro Phe Ser  
 65 70 75 80  
 Gln Gln Gln Pro Ile Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln  
 85 90 95  
 Gln Leu Val Leu Pro Gln Gln Pro Pro Phe Ser Gln Gln Gln Gln Pro  
 100 105 110  
 Val Leu Pro Pro Gln Gln Ser Pro Phe Pro Gln Gln Gln Gln Gln His  
 115 120 125  
 Gln Gln Leu Val Gln Gln Gln Ile Pro Val Val Gln Pro Ser Ile Leu  
 130 135 140  
 Gln Gln Leu Asn Pro Cys Lys Val Phe Leu Gln Gln Gln Cys Ser Pro  
 145 150 155 160  
 Val Ala Met Pro Gln Arg Leu Ala Arg Ser Gln Met Leu Gln Gln Ser  
 165 170 175  
 Ser Cys His Val Met Gln Gln Gln Cys Cys Gln Gln Leu Pro Gln Ile  
 180 185 190

## 116

Pro Gln Gln Ser Arg Tyr Glu Ala Ile Arg Ala Ile Ile Tyr Ser Ile  
 195 200 205

Ile Leu Gln Glu Gln Gln Gln Val Gln Gly Ser Ile Gln Ser Gln Gln  
 210 215 220

Gln Gln Pro Gln Gln Leu Gly Gln Cys Val Ser Gln Pro Gln Gln Gln  
 225 230 235 240

Ser Gln Gln Gln Leu Gly Gln Gln Pro Gln Gln Gln Gln Leu Ala Gln  
 245 250 255

Gly Thr Phe Leu Gln Pro His Gln Ile Ala Gln Leu Glu Val Met Thr  
 260 265 270

Ser Ile Ala Leu Arg Ile Leu Pro Thr Met Cys Ser Val Asn Val Pro  
 275 280 285

Leu Tyr Arg Thr Thr Thr Ser Val Pro Phe Gly Val Gly Thr Gly Val  
 290 295 300

Gly Ala Tyr  
 305

&lt;210&gt; 114

&lt;211&gt; 8482

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: binary  
 expression vector

&lt;400&gt; 114

```

ttccatggac atacaaatgg acgaacggat aaaccttttc acgccctttt aaatatccga 60
ttatttctaataaacgctctt ttctcttagg tttacccgcc aatatatcct gtcaaact 120
gatagtttaa actgaaggcg ggaaacgaca atcagatcta gtaggaacaa gctatgacca 180
tgattacgcc aatcaccact ttgtacaaga aagctgggtc tagatgacgg acaatcagta 240
aattgaacgg agaataattat tcataaaaaat acgatagtaa cgggtgatat attcattaga 300
atgaaccgaa accggcggtta aggatctgag ctacacatgc tcagggtttt tacaacgtgc 360
acaacagaat tgaaagcaaa tatcatgcga tcataggcgt ctgcataatc tcattaaagc 420
aggaggcctt ctagactgca ggcgccgcc caccgcgggtg ggctggctat gaagaaatta 480
taatcgtgta aaacttagtg agtgtgtatg aatgaaagta ttgcaaaatc ctcattatat 540
agactacatg cataactagt tgcattgtaa tttgtagttt tcttcattat tgcatectcc 600
aagtggatgt catgggttta cacatggctt ccatgcaaat catttccaaa atatttttaa 660
actttccaca gggcatccat gcatgcacct caaaacttgt gtgtggtaac attgtgtct 720
tgaaaaatta ctaaaccctt tgtccacgtg acgttcatgc acctcaaate ttgtgtggta 780
ccattattat cctcaagaat tattgaatgt ttggtgtata tgccatccat gcagcattgc 840
aacaattaaa tctccaaacc ttgtggtacc atattcactc actttaattc tcctatagta 900
gaaatattag caaatattta catttccagt tgattagtat atgtatttag aagacaaaaa 960
taatttagaa tcaattaatc aacttgcaaa ttgctaagtg ttggcaaacg ttagcataaa 1020
agtggttata aatttagtac caaatataaa aatttatcgc aaatcaaata cataacacac 1080
atagtaaaac aaaaacaaat tacaagggtt tagacgttta gtggcaatgt gtaaatattgc 1140
tcgactgaat tggttccttt aagcctgctt tttgtacaa acttgtgata attcactggc 1200
cgtcgtttta caacgactca ggatcctgtc aaacactgat agtttaaact gaaggcggga 1260
aacgacaatc tgatcatgag cggagaatta agggagtcac gttatgaccc ccgccgatga 1320
cgcgggacaa gccgttttac gtttggaact gacagaaccg caacgttgaa ggagccactc 1380
agccgcgggt ttctggagtt taatgagcta agcacatacg tcagaaacca ttattgcgcg 1440
ttcaaaagtc gcctaaggtc actatcagct agcaaatatt tcttgtcaaa aatgctccac 1500
tgacgttcca taaattcccc tcggtatcca attagagtct catattcact ctcaatccaa 1560
ataatctgca ccggatctgg atcgtttcgc atgattgaac aagatggatt gcacgcaggt 1620
tctccggccg cttgggtgga gaggctattc ggctatgact gggcacaaca gacaatcggc 1680

```

tgctctgatg	ccgccgtggt	ccggctgtca	gcgcaggggc	gccccggttct	ttttgtcaag	1740
accgacctgt	ccggtgccct	gaatgaactg	caggacgagg	cagcgcgggt	atcgtggctg	1800
gccacgacgg	geggttcctt	cgcagctgtg	ctcgacgttg	tcactgaagc	gggaagggac	1860
tggctgctat	tgggcgaagt	gccggggcag	gatctcctgt	catctcacct	tgctcctgcc	1920
gagaaagtat	ccatcatggc	tgatgcaatg	cggcggtctg	atacgcttga	tccggctacc	1980
tgcccatctg	accaccaagc	gaaacatcgc	atcgagcgag	cacgtactcg	gatggaagcc	2040
ggtcttgctg	atcaggatga	tctggacgaa	gagcatcagg	ggctcgcgcc	agccgaactg	2100
ttcgccaggc	tcaaggcgcg	catgcccgcg	ggcgaggatc	tcgctcgtgac	ccatggcgat	2160
gcctgcttgc	cgaatatcat	ggtggaaaat	ggccgctttt	ctggattcat	cgactgtggc	2220
cggctgggtg	tggcggaacc	ctatcaggac	atagcgttgg	ctacccgtga	tattgctgaa	2280
gagcttggcg	gcgaatgggc	tgacogcttc	ctcgtgcttt	acggtatcgc	cgctcccgat	2340
tcgcagcgca	tcgccttcta	tcgccttctt	gacgagttct	tctgagcggg	accecaagctc	2400
tagatcttgc	tgcgttcgga	tattttcgtg	gagttcccgc	cacagacccg	gatgatcccc	2460
gatcgttcaa	acatttggca	ataaagtttc	ttaagattga	atcctgttgc	cggctcttgcg	2520
atgattatca	tataatttct	gttgaattac	gttaagcatg	taataattaa	catgtaatgc	2580
atgacgttat	ttatgagatg	ggtttttatg	attagagtcc	cgcaattata	catttaatac	2640
gcgatagaaa	acaaaatata	gcgcgcaaac	taggataaat	tatcgcgcgc	ggtgtcatct	2700
atgttactag	atcgggcctc	ctgtcaagct	ctgcttggtg	ataattgtca	ttagattggt	2760
tttatgcata	gatgcactcg	aaatcagcca	atttttagaca	agtatcaaac	ggatgttaat	2820
tcagtacatt	aaagacgtcc	gcaatgtgtt	attaagttgt	ctaagcgtca	atttgtttac	2880
accacaatat	atcctgccac	cagccagcca	acagctcccc	gaccggcagc	tccgcacaaa	2940
atcaccacgc	gttaccacca	cgcgggcggg	ccgcattggt	ttgaccgtgt	tcgcccggcat	3000
tgccgagttc	gagcgttccc	taatcatcga	cgcgacccgg	agcgggcggc	aggccgccaa	3060
ggcccgaggc	gtgaagtttg	gcccccgccc	taccctcacc	ccggcacaga	tcgcgcacgc	3120
ccgcgagctg	atcgaccagg	aaggccgcac	cgtgaaagag	gcggctgcac	tgcttggcgt	3180
gcacgcctcg	accctgtacc	gcgcacttga	gcgcagcgag	gaagtgcacg	ccaccgggac	3240
caggcgggcg	ggtgccttcc	gtgaggacgc	attgaccgag	gccgacgccc	tggcgggccg	3300
cgagaatgaa	cgccaagagg	aacaagcatg	aaaccgcacc	aggacggcca	ggacgaaccg	3360
tttttcatta	ccgaagagat	cgaggcggag	atgatcgcgg	ccgggtacgt	gttcgagccg	3420
cccgcgcacg	tctcaaccgt	gcggctgcat	gaaatcctgg	ccggtttgtc	tgatgccaa	3480
ctggcgggcct	ggccggccag	cttggccgct	gaagaaaccg	agcggccgccc	tctaaaaagg	3540
tgatgtgtat	ttgagtaaaa	cagcttgctg	catgcggctg	ctgcgtatat	gatgcgatga	3600
gtaataaac	aaatacgcaa	ggggaacgca	tgaaggttat	cgctgtactt	aaccagaaag	3660
gcgggtcagg	caagacgacc	atcgcaaccc	atctagcccg	cgccctgcaa	ctcgccgggg	3720
ccgatgttct	gttagtcgat	tccgatcccc	agggcagtg	ccgcgattgg	gcggccgtgc	3780
gggaagatca	accgctaacc	gttgctcgga	tcgaccgccc	gacgattgac	cgcgacgtga	3840
aggccatcgg	ccggcgcgac	ttcgtagtga	tcgacggagc	gccccaggcg	gcggacttgg	3900
ctgtgtccgc	gatcaaggca	gccgacttgc	tgctgattcc	ggtgcagcca	agcccttacg	3960
acatatgggc	caccgccgac	ctggtggagc	tgggttaagca	gcgcattgag	gtcacggatg	4020
gaaggctaca	agcggccttt	gtcgtgtcgc	gggcgatcaa	aggcacgcgc	atcggcggtg	4080
agggtgccga	ggcgttggcc	gggtacgagc	tgccatttct	tgagtcccgt	atcacgcagc	4140
gcgtgagcta	cccaggcact	gccgcgcggc	gcacaaccgt	tcttgaatca	gaacccgagg	4200
gcgacgctgc	ccgcgaggtc	caggcgctgg	ccgctgaaat	taaatcaaaa	ctcatttgag	4260
ttaatgaggt	aaagagaaaa	tgagcaaaa	cacaaacacg	ctaagtgccg	gccgtccgag	4320
cgcacgcagc	agcaaggctg	caacgttggc	cagcctggca	gacacgccag	ccatgaagcg	4380
ggtcaacttt	cagttgccgg	cggaggatca	caccaagctg	aagatgtacg	cggtagccca	4440
aggcaagacc	attaccgagc	tgctatctga	atacatcgcg	cagctaccag	agtaaatgag	4500
caaatgaata	aatgagtaga	tgaatttttag	cggctaaagg	aggcggcatg	gaaaatcaag	4560
aacaaccagg	caccgacgcc	gtggaatgcc	ccatgtgtgg	aggaacgggc	ggttggccag	4620
gcgtaagcgg	ctgggttgtc	tgccggccct	gcaatggcac	tggaaacccc	aagcccagg	4680
aatcggcgtg	agcggtcgca	aaccatccgg	cccgttaca	atcggcgcg	cgctgggtga	4740
tgacctggtg	gagaagttga	aggccgcgca	ggccgcccag	cggcaacgca	tcgaggcaga	4800
agcacgcccc	ggtgaatcgt	ggcaagcggc	cgctgatcga	atccgcaaag	aatcccggca	4860
accgcccggca	gccggtgcgc	cgctcattag	gaagccgccc	aagggcgacg	agcaaccaga	4920
ttttttcgtt	ccgatgctct	atgacgtggg	caccgcgat	agtcgcagca	tcattggacgt	4980
ggccgttttc	cgtctgtcga	agcgtgaccg	acgagctggc	gaggtgatcc	gctacgagct	5040
tccagacggg	cacgtagagg	tttcgcgagg	gccggccggc	atggccagtg	tgtgggatta	5100

cgacctggta	ctgatggcgg	tttcccatct	aaccgaatcc	atgaaccgat	accgggaagg	5160
gaagggagac	aagcccggcc	gcgtgttccg	tccacacgtt	gcggacgtac	tcaagttctg	5220
ccggcgagcc	gatggcggaa	agcagaaaga	cgacctggta	gaaacctgca	ttcgggttaa	5280
caccacgcac	gttgccatgc	agcgtacgaa	gaaggccaag	aacggccgcc	tggtgacggt	5340
atccgagggg	gaagccttga	ttagccgcta	caagatcgta	aagagcgaaa	ccgggcccgc	5400
ggagtacatc	gagatcgagc	tagctgattg	gatgtaccgc	gagatcacag	aaggcaagaa	5460
cccggacgtg	ctgacggttc	accccgatta	ctttttgata	gatcccggca	tcggccggtt	5520
tctctaccgc	ctggcacgcc	gcgccgcagg	caaggcagaa	gccagatggt	tggtcaagac	5580
gatctacgaa	cgcagtggca	gcgccggaga	gttcaagaag	ttctgtttca	ccgtgcgcaa	5640
gctgatcggg	tcaaatgacc	tgccggagta	cgatttgaag	gaggaggcgg	ggcaggctgg	5700
cccgatcccta	gtcatgcgct	accgcaacct	gatcgagggc	gaagcatccg	ccggttccta	5760
atgtacggag	cagatgctag	ggcaaattgc	cctagcaggg	gaaaaaggtc	gaaaagggtc	5820
ctttcctgtg	gatagcacgt	acattgggaa	cccaaagccg	tacattggga	accggaacct	5880
gtacattggg	aacccaaagc	cgtacattgg	gaaccgggtc	cacatgtaag	tgactgatat	5940
aaaagagaaa	aaaggcgatt	tttccgccta	aaactcttta	aaacttatta	aaactcttaa	6000
aaccgcctg	gcctgtgcat	aactgtctgg	ccagcgcaca	gccgaagagc	tgcaaaaagc	6060
gcctaccctt	cggtcgctgc	gctccctacg	ccccgcgcgt	tcgcgtcggc	ctatcgcggc	6120
cgctggccgc	tcaaaaatgg	ctggcctacg	gccaggcaat	ctaccagggc	gcggacaagc	6180
cgcgccgtcg	ccactcgacc	gccggcgccc	acatcaaggc	accctgcctc	gcgcgtttcg	6240
gtgatgacgg	tgaaaacctc	tgacacatgc	agctcccggg	gacggtcaca	gcttgtctgt	6300
aagcggatgc	cgggagcaga	caagcccgtc	agggcgcgct	agcgggtggt	ggcgggtgtc	6360
ggggcgagc	catgaccag	tcacgtagcg	atagcggagt	gtatactggc	ttaactatgc	6420
ggcatcagag	cagattgtac	tgagagtgca	ccatatgcgg	tgtgaaatac	cgcacagatg	6480
cgtaaggaga	aaataccgca	tcaggcgctc	ttccgcttcc	tcgctcactg	actcgtctgc	6540
ctcggtcggt	cggctgcggc	gagcgggtatc	agctcactca	aaggcggtaa	tacggttatc	6600
cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	aaaaggccag	6660
gaaccgtaaa	aaggccgcgt	tgctggcggt	tttccatagg	ctccgcccc	ctgacgagca	6720
tcacaaaaat	cgacgctcaa	gtcagagggtg	gcgaaaccgg	acaggactat	aaagatacca	6780
ggcggtttccc	cctggaagct	ccctcggtgc	ctctcctggt	ccgaccctgc	cgcttaccgg	6840
atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	tctcatagct	cacgctgtag	6900
gtatctcagt	tcggtgtagg	tcgttcgctc	caagctgggc	tgtgtgcacg	aaccccccg	6960
tcagcccagc	cgctgcgcct	tatccggtaa	ctatcgtctt	gagtcacaac	cggtaagaca	7020
cgacttatcg	ccactggcag	cagccactgg	taacaggatt	agcagagcga	ggtatgtagg	7080
cgggtgtaca	gagttcttga	agtgggtggc	taactacggc	tacactagaa	ggacagtatt	7140
tggtatctgc	gctctgctga	agccagttac	cttcggaaaa	agagttggta	gctcttgatc	7200
cggcaaaaca	accaccgctg	gtagcgggtg	ttttttgtt	tgcaagcagc	agattacgcg	7260
cagaaaaaaa	ggatctcaag	aagatccttt	gatcttttct	acggggtctg	acgctcagtg	7320
gaacgaaaaa	tcacgttaag	ggattttggg	catgcatgat	atatctccca	atttgtgtag	7380
ggettattat	gcacgcttaa	aaataataaa	agcagacttg	acctgatagt	ttggctgtga	7440
gcaattatgt	gcttagtgca	tctaacgctt	gagttaaagg	gcgcgcgcaa	gcggcgctcg	7500
cttgaacgaa	tttctagcta	gacattatct	gccgactacc	ttgggtgatct	cgcctttcac	7560
gtagtggaca	aattcttcca	actgatctgc	gcgcgaggcc	aagcgatctt	cttcttgtcc	7620
aagataagcc	tgtctagctt	caagtatgac	gggctgatac	tgggccggca	ggcgctccat	7680
tgcccagtcg	gcagcgacat	ccttcggcgc	gattttgcgc	gttactgcgc	tgtaccaa	7740
gcgggacaac	gtaagcacta	catttcgctc	atcgccagcc	cagtcgggcg	gcgagttcca	7800
tagcgttaag	gtttcattta	gcgcctcaaa	tagatcctgt	tcaggaaccg	gatcaaagag	7860
ttcctccgcc	gctggaccta	ccaaggcaac	gctatgttct	cttgcttttg	tcagcaagat	7920
agccagatca	atgtcgatcg	tggttggtc	gaagatacct	gcaagaatgt	cattgcgctg	7980
ccattctcca	aattgcagtt	cgcgcttagc	tggtataacgc	cacggaatga	tgctcgctcg	8040
cacaacaatg	gtgacttcta	cagcgcggag	aatctcgctc	tctccagggg	aagccgaagt	8100
ttccaaaagg	tcgttgatca	aagctcgccg	cgttggttca	tcaagcctta	cggtcaccgt	8160
aaccagcaaa	tcaatatcac	tgtgtggctt	caggccgcca	tccactgcgg	agccgtacaa	8220
atgtacggcc	agcaacgtcg	gttcgagatg	gcgctcgatg	acgccaaacta	cctctgatag	8280
ttgagtcgat	acttcggcga	tcaccgcttc	ccccatgatg	tttaactttg	ttttagggcg	8340
actgccttgc	tgcgtaacat	cgttgctgct	ccataacatc	aaacatcgac	ccacggcgta	8400
acgcgcttgc	tgcttggtatg	cccgaggcat	agactgtacc	ccaaaaaaac	agtcataaca	8460
agccatgaaa	accgccactg	cg				8482

119

```

<210> 115
<211> 575
<212> DNA
<213> Brassica napus
<220>
<221> misc_feature
<222> (1)..(6)
<223> restriction site
<220>
<221> misc_feature
<222> (570)..(575)
<223> restriction site
<220>
<221> misc_feature
<222> (7)..(575)
<223> coding for homogentisate-1,2-dioxygenase (HDG)
<400> 115
gtcgacgggc cgatgggggc gaagggtctt gctgcaccaa gagattttct tgcaccaacg 60
gcatggtttg aggaagggtt acggcctgac tacactattg ttcagaagtt tggcggtgaa 120
ctctttactg ctaaacaaga tttctctccg ttcaatgttg ttgcctggca tggcaattac 180
gtgccttata agtatgacct gcacaagttc tgtccataca aactgtcct tgtagaccat 240
ggagatccat ctgtaaatac agttctgaca gcaccaacgg ataaacctgg tgtggccttg 300
cttgattttg tcatattccc tcctcgttgg ttggttgctg agcatacctt tcgacctcct 360
tactaccatc gtaactgcat gagtgaattt atgggcctaa tctatggtgc ttacgaggcc 420
aaagctgatg gatttctacc tggtgggcga agtcttcaca gttgtatgac acctcatggt 480
ccagatacaa ccacatacga ggcgacgatt gctcgtgtaa atgcaatggc tccttataag 540
ctcacaggca ccatggcctt catgtttgag gtacc 575

<210> 116
<211> 1386
<212> DNA
<213> Arabidopsis thaliana
<220>
<221> CDS
<222> (1)..(1383)
<223> coding for homogentisate-1,2-dioxygenase (HDG)
<400> 116
atg gaa gag aag aag aag gag ctt gaa gag ttg aag tat caa tca ggt 48
Met Glu Glu Lys Lys Lys Glu Leu Glu Glu Leu Lys Tyr Gln Ser Gly
  1             5             10             15

ttt ggt aac cac ttc tca tcg gaa gca atc gcc gga gct tta ccg tta 96
Phe Gly Asn His Phe Ser Ser Glu Ala Ile Ala Gly Ala Leu Pro Leu
      20             25             30

gat cag aac agt cct ctt ctt tgt cct tac ggt ctt tac gcc gaa cag 144
Asp Gln Asn Ser Pro Leu Leu Cys Pro Tyr Gly Leu Tyr Ala Glu Gln
      35             40             45

atc tcc ggt act tct ttc act tct cct cgc aag ctc aat caa aga agt 192
Ile Ser Gly Thr Ser Phe Thr Ser Pro Arg Lys Leu Asn Gln Arg Ser
      50             55             60

tgg ttg tac cgg gtt aaa cca tcg gtt aca cat gaa ccg ttc aag cct 240
Trp Leu Tyr Arg Val Lys Pro Ser Val Thr His Glu Pro Phe Lys Pro
      65             70             75             80

cgt gta cca gct cat aag aag ctt gtg agt gag ttt gat gca tca aat 288
Arg Val Pro Ala His Lys Lys Leu Val Ser Glu Phe Asp Ala Ser Asn
      85             90             95

```

## 120

agt	cgt	acg	aat	ccg	act	cag	ctt	cgg	tgg	aga	cct	gag	gat	att	cct	336
Ser	Arg	Thr	Asn	Pro	Thr	Gln	Leu	Arg	Trp	Arg	Pro	Glu	Asp	Ile	Pro	
			100					105						110		
gat	tcg	gag	att	gat	ttc	gtt	gat	ggg	tta	ttt	acc	att	tgt	gga	gct	384
Asp	Ser	Glu	Ile	Asp	Phe	Val	Asp	Gly	Leu	Phe	Thr	Ile	Cys	Gly	Ala	
		115					120					125				
gga	agc	tcg	ttt	ctt	cgc	cat	ggc	ttc	gct	att	cac	atg	tat	gtg	gct	432
Gly	Ser	Ser	Phe	Leu	Arg	His	Gly	Phe	Ala	Ile	His	Met	Tyr	Val	Ala	
		130				135					140					
aac	aca	gga	atg	aaa	gac	tcc	gca	ttt	tgc	aac	gct	gat	ggg	gac	ttc	480
Asn	Thr	Gly	Met	Lys	Asp	Ser	Ala	Phe	Cys	Asn	Ala	Asp	Gly	Asp	Phe	
145					150					155					160	
ttg	tta	gtt	cct	caa	aca	gga	agg	cta	tgg	att	gaa	act	gag	tgt	gga	528
Leu	Leu	Val	Pro	Gln	Thr	Gly	Arg	Leu	Trp	Ile	Glu	Thr	Glu	Cys	Gly	
				165					170					175		
agg	ctt	ttg	gta	act	cct	ggt	gag	att	gct	gtt	ata	cca	caa	ggt	ttc	576
Arg	Leu	Leu	Val	Thr	Pro	Gly	Glu	Ile	Ala	Val	Ile	Pro	Gln	Gly	Phe	
			180					185					190			
cgt	ttc	tcc	ata	gat	tta	ccg	gat	ggg	aag	tct	cgt	ggt	tat	gtt	gct	624
Arg	Phe	Ser	Ile	Asp	Leu	Pro	Asp	Gly	Lys	Ser	Arg	Gly	Tyr	Val	Ala	
		195					200					205				
gaa	atc	tat	ggg	gct	cat	ttt	cag	ctt	cct	gat	ctt	gga	cca	ata	ggt	672
Glu	Ile	Tyr	Gly	Ala	His	Phe	Gln	Leu	Pro	Asp	Leu	Gly	Pro	Ile	Gly	
	210					215					220					
gct	aat	ggt	ctt	gct	gca	tca	aga	gat	ttt	ctt	gca	cca	aca	gca	tgg	720
Ala	Asn	Gly	Leu	Ala	Ala	Ser	Arg	Asp	Phe	Leu	Ala	Pro	Thr	Ala	Trp	
225					230					235					240	
ttt	gag	gat	gga	ttg	cgg	cct	gaa	tac	aca	att	gtt	cag	aag	ttt	ggc	768
Phe	Glu	Asp	Gly	Leu	Arg	Pro	Glu	Tyr	Thr	Ile	Val	Gln	Lys	Phe	Gly	
			245					250						255		
ggt	gaa	ctc	ttt	act	gct	aaa	caa	gat	ttc	tct	cca	ttc	aat	gtg	gtt	816
Gly	Glu	Leu	Phe	Thr	Ala	Lys	Gln	Asp	Phe	Ser	Pro	Phe	Asn	Val	Val	
			260					265					270			
gcc	tgg	cat	ggc	aat	tac	gtg	cct	tat	aag	tat	gac	ctg	aag	aag	ttc	864
Ala	Trp	His	Gly	Asn	Tyr	Val	Pro	Tyr	Lys	Tyr	Asp	Leu	Lys	Lys	Phe	
		275					280					285				
tgt	cca	tac	aac	act	gtg	ctt	tta	gat	cat	gga	gat	cca	tct	ata	aat	912
Cys	Pro	Tyr	Asn	Thr	Val	Leu	Leu	Asp	His	Gly	Asp	Pro	Ser	Ile	Asn	
	290					295					300					
aca	gtc	ctt	aca	gca	cca	act	gat	aaa	cct	ggt	gtg	gcc	ttg	ctt	gat	960
Thr	Val	Leu	Thr	Ala	Pro	Thr	Asp	Lys	Pro	Gly	Val	Ala	Leu	Leu	Asp	
305					310					315					320	
ttt	gtc	ata	ttt	cct	cct	cga	tgg	ttg	gtt	gct	gag	cat	act	ttt	cga	1008
Phe	Val	Ile	Phe	Pro	Pro	Arg	Trp	Leu	Val	Ala	Glu	His	Thr	Phe	Arg	
				325					330					335		
cct	cct	tac	tat	cat	cgt	aac	tgc	atg	agt	gaa	ttt	atg	ggc	tta	atc	1056
Pro	Pro	Tyr	Tyr	His	Arg	Asn	Cys	Met	Ser	Glu	Phe	Met	Gly	Leu	Ile	
			340				345						350			
tac	ggt	gca	tac	gag	gcg	aaa	gct	gat	gga	ttt	ctc	cct	ggc	ggt	gca	1104
Tyr	Gly	Ala	Tyr	Glu	Ala	Lys	Ala	Asp	Gly	Phe	Leu	Pro	Gly	Gly	Ala	
		355					360						365			

## 121

```

agt ctt cat agc tgt atg aca cct cat ggt cca gat act acc acg tac 1152
Ser Leu His Ser Cys Met Thr Pro His Gly Pro Asp Thr Thr Thr Tyr
    370                      375                      380

gag gcg aca att gct cga gta aat gca atg gct cct tct aaa ctc aca 1200
Glu Ala Thr Ile Ala Arg Val Asn Ala Met Ala Pro Ser Lys Leu Thr
    385                      390                      395                      400

ggt acg atg gct ttc atg ttc gaa tca gca ttg atc cct aga gtc tgt 1248
Gly Thr Met Ala Phe Met Phe Glu Ser Ala Leu Ile Pro Arg Val Cys
                      405                      410                      415

cat tgg gct ctg gag tct cct ttc ctg gat cac gac tac tac cag tgt 1296
His Trp Ala Leu Glu Ser Pro Phe Leu Asp His Asp Tyr Tyr Gln Cys
                      420                      425                      430

tgg att ggc ctc aag tct cat ttc tcg cgc ata agc ttg gac aag aca 1344
Trp Ile Gly Leu Lys Ser His Phe Ser Arg Ile Ser Leu Asp Lys Thr
    435                      440                      445

aat gtt gaa tca aca gag aaa gaa cca gga gct tcg gag taa 1386
Asn Val Glu Ser Thr Glu Lys Glu Pro Gly Ala Ser Glu
    450                      455                      460

<210> 117
<211> 461
<212> PRT
<213> Arabidopsis thaliana

<400> 117
Met Glu Glu Lys Lys Lys Glu Leu Glu Glu Leu Lys Tyr Gln Ser Gly
    1                      5                      10                      15
Phe Gly Asn His Phe Ser Ser Glu Ala Ile Ala Gly Ala Leu Pro Leu
    20                      25                      30
Asp Gln Asn Ser Pro Leu Leu Cys Pro Tyr Gly Leu Tyr Ala Glu Gln
    35                      40                      45
Ile Ser Gly Thr Ser Phe Thr Ser Pro Arg Lys Leu Asn Gln Arg Ser
    50                      55                      60
Trp Leu Tyr Arg Val Lys Pro Ser Val Thr His Glu Pro Phe Lys Pro
    65                      70                      75                      80
Arg Val Pro Ala His Lys Lys Leu Val Ser Glu Phe Asp Ala Ser Asn
    85                      90                      95
Ser Arg Thr Asn Pro Thr Gln Leu Arg Trp Arg Pro Glu Asp Ile Pro
    100                      105                      110
Asp Ser Glu Ile Asp Phe Val Asp Gly Leu Phe Thr Ile Cys Gly Ala
    115                      120                      125
Gly Ser Ser Phe Leu Arg His Gly Phe Ala Ile His Met Tyr Val Ala
    130                      135                      140
Asn Thr Gly Met Lys Asp Ser Ala Phe Cys Asn Ala Asp Gly Asp Phe
    145                      150                      155                      160
Leu Leu Val Pro Gln Thr Gly Arg Leu Trp Ile Glu Thr Glu Cys Gly
    165                      170                      175
Arg Leu Leu Val Thr Pro Gly Glu Ile Ala Val Ile Pro Gln Gly Phe
    180                      185                      190
Arg Phe Ser Ile Asp Leu Pro Asp Gly Lys Ser Arg Gly Tyr Val Ala
    195                      200                      205

```



## 122

Glu Ile Tyr Gly Ala His Phe Gln Leu Pro Asp Leu Gly Pro Ile Gly  
 210 215 220  
 Ala Asn Gly Leu Ala Ala Ser Arg Asp Phe Leu Ala Pro Thr Ala Trp  
 225 230 235 240  
 Phe Glu Asp Gly Leu Arg Pro Glu Tyr Thr Ile Val Gln Lys Phe Gly  
 245 250 255  
 Gly Glu Leu Phe Thr Ala Lys Gln Asp Phe Ser Pro Phe Asn Val Val  
 260 265 270  
 Ala Trp His Gly Asn Tyr Val Pro Tyr Lys Tyr Asp Leu Lys Lys Phe  
 275 280 285  
 Cys Pro Tyr Asn Thr Val Leu Leu Asp His Gly Asp Pro Ser Ile Asn  
 290 295 300  
 Thr Val Leu Thr Ala Pro Thr Asp Lys Pro Gly Val Ala Leu Leu Asp  
 305 310 315 320  
 Phe Val Ile Phe Pro Pro Arg Trp Leu Val Ala Glu His Thr Phe Arg  
 325 330 335  
 Pro Pro Tyr Tyr His Arg Asn Cys Met Ser Glu Phe Met Gly Leu Ile  
 340 345 350  
 Tyr Gly Ala Tyr Glu Ala Lys Ala Asp Gly Phe Leu Pro Gly Gly Ala  
 355 360 365  
 Ser Leu His Ser Cys Met Thr Pro His Gly Pro Asp Thr Thr Thr Tyr  
 370 375 380  
 Glu Ala Thr Ile Ala Arg Val Asn Ala Met Ala Pro Ser Lys Leu Thr  
 385 390 395 400  
 Gly Thr Met Ala Phe Met Phe Glu Ser Ala Leu Ile Pro Arg Val Cys  
 405 410 415  
 His Trp Ala Leu Glu Ser Pro Phe Leu Asp His Asp Tyr Tyr Gln Cys  
 420 425 430  
 Trp Ile Gly Leu Lys Ser His Phe Ser Arg Ile Ser Leu Asp Lys Thr  
 435 440 445  
 Asn Val Glu Ser Thr Glu Lys Glu Pro Gly Ala Ser Glu  
 450 455 460

&lt;210&gt; 118

&lt;211&gt; 815

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (37)..(705)

&lt;223&gt; coding for maleylacetoacetate isomerase (MAAI)

&lt;400&gt; 118

gtaatctccg aagaagaaca aattccttgc tgaatc atg tct tat gtt acc gat 54  
 Met Ser Tyr Val Thr Asp  
 1 5

ttt tat cag gcg aag ttg aag ctc tac tct tac tgg aga agc tca tgt 102  
 Phe Tyr Gln Ala Lys Leu Lys Leu Tyr Ser Tyr Trp Arg Ser Ser Cys  
 10 15 20

## 123

gct cat cgc gtc cgt atc gcc ctc act tta aaa ggg ctt gat tat gaa 150  
 Ala His Arg Val Arg Ile Ala Leu Thr Leu Lys Gly Leu Asp Tyr Glu  
                   25                                  30                                  35

tat ata ccg gtt aat ttg ctc aaa ggg gat caa tcc gat tca gat ttc 198  
 Tyr Ile Pro Val Asn Leu Leu Lys Gly Asp Gln Ser Asp Ser Asp Phe  
                   40                                  45                                  50

aag aag atc aat cca atg ggc act gta cca gcg ctt gtt gat ggt gat 246  
 Lys Lys Ile Asn Pro Met Gly Thr Val Pro Ala Leu Val Asp Gly Asp  
                   55                                  60                                  65                                  70

gtt gtg att aat gac tct ttc gca ata ata atg tac ctg gat gat aag 294  
 Val Val Ile Asn Asp Ser Phe Ala Ile Ile Met Tyr Leu Asp Asp Lys  
                                   75                                  80                                  85

tat ccg gag cca ccg ctg tta cca agt gac tac cat aaa cgg gcg gta 342  
 Tyr Pro Glu Pro Pro Leu Leu Pro Ser Asp Tyr His Lys Arg Ala Val  
                                   90                                  95                                  100

aat tac cag gcg acg agt att gtc atg tct ggt ata cag cct cat caa 390  
 Asn Tyr Gln Ala Thr Ser Ile Val Met Ser Gly Ile Gln Pro His Gln  
                   105                                  110                                  115

aat atg gct ctt ttt agg tat ctc gag gac aag ata aat gct gag gag 438  
 Asn Met Ala Leu Phe Arg Tyr Leu Glu Asp Lys Ile Asn Ala Glu Glu  
                   120                                  125                                  130

aaa act gct tgg att act aat gct atc aca aaa gga ttc aca gct ctc 486  
 Lys Thr Ala Trp Ile Thr Asn Ala Ile Thr Lys Gly Phe Thr Ala Leu  
                   135                                  140                                  145                                  150

gag aaa ctg ttg gtg agt tgc gct gga aaa tac gcg act ggt gat gaa 534  
 Glu Lys Leu Leu Val Ser Cys Ala Gly Lys Tyr Ala Thr Gly Asp Glu  
                                   155                                  160                                  165

gtt tac ttg gct gat ctt ttc cta gca cca cag atc cac gca gca ttc 582  
 Val Tyr Leu Ala Asp Leu Phe Leu Ala Pro Gln Ile His Ala Ala Phe  
                                   170                                  175                                  180

aac aga ttc cat att aac atg gaa cca ttc ccg act ctt gca agg ttt 630  
 Asn Arg Phe His Ile Asn Met Glu Pro Phe Pro Thr Leu Ala Arg Phe  
                   185                                  190                                  195

tac gag tca tac aac gaa ctg cct gca ttt caa aat gca gtc ccg gag 678  
 Tyr Glu Ser Tyr Asn Glu Leu Pro Ala Phe Gln Asn Ala Val Pro Glu  
                   200                                  205                                  210

aag caa cca gat act cct tcc acc atc tgattctgtg aaccgtaagc 725  
 Lys Gln Pro Asp Thr Pro Ser Thr Ile  
                   215                                  220

ttctctcagt ctcagctcaa taaaatctct taggaaacaa caacaacacc ttgaacttaa 785  
 atgtatcata tgaaccagtt tacaataat 815

<210> 119  
 <211> 223  
 <212> PRT  
 <213> Arabidopsis thaliana  
 <400> 119  
 Met Ser Tyr Val Thr Asp Phe Tyr Gln Ala Lys Leu Lys Leu Tyr Ser  
                   1                                  5                                  10                                  15  
 Tyr Trp Arg Ser Ser Cys Ala His Arg Val Arg Ile Ala Leu Thr Leu  
                                   20                                  25                                  30

## 124

Lys Gly Leu Asp Tyr Glu Tyr Ile Pro Val Asn Leu Leu Lys Gly Asp  
                   35                                  40                                  45  
 Gln Ser Asp Ser Asp Phe Lys Lys Ile Asn Pro Met Gly Thr Val Pro  
           50                                  55                                  60  
 Ala Leu Val Asp Gly Asp Val Val Ile Asn Asp Ser Phe Ala Ile Ile  
   65                                  70                                  75                                  80  
 Met Tyr Leu Asp Asp Lys Tyr Pro Glu Pro Pro Leu Leu Pro Ser Asp  
                                   85                                  90                                  95  
 Tyr His Lys Arg Ala Val Asn Tyr Gln Ala Thr Ser Ile Val Met Ser  
                                   100                                  105                                  110  
 Gly Ile Gln Pro His Gln Asn Met Ala Leu Phe Arg Tyr Leu Glu Asp  
                   115                                  120                                  125  
 Lys Ile Asn Ala Glu Glu Lys Thr Ala Trp Ile Thr Asn Ala Ile Thr  
   130                                  135                                  140  
 Lys Gly Phe Thr Ala Leu Glu Lys Leu Leu Val Ser Cys Ala Gly Lys  
  145                                  150                                  155                                  160  
 Tyr Ala Thr Gly Asp Glu Val Tyr Leu Ala Asp Leu Phe Leu Ala Pro  
                                   165                                  170                                  175  
 Gln Ile His Ala Ala Phe Asn Arg Phe His Ile Asn Met Glu Pro Phe  
                                   180                                  185                                  190  
 Pro Thr Leu Ala Arg Phe Tyr Glu Ser Tyr Asn Glu Leu Pro Ala Phe  
                   195                                  200                                  205  
 Gln Asn Ala Val Pro Glu Lys Gln Pro Asp Thr Pro Ser Thr Ile  
   210                                  215                                  220

&lt;210&gt; 120

&lt;211&gt; 1227

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis thaliana

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(1224)

&lt;223&gt; coding for fumarylacetoacetate hydrolase (FAAH)

&lt;400&gt; 120

atg gcg ttg ctg aag tct ttc atc gat gtt ggc tca gac tcg cac ttc 48  
 Met Ala Leu Leu Lys Ser Phe Ile Asp Val Gly Ser Asp Ser His Phe  
   1                                  5                                  10                                  15

cct atc cag aat ctg cct tat ggt gtc ttc aaa ccg gaa tcg aac tca 96  
 Pro Ile Gln Asn Leu Pro Tyr Gly Val Phe Lys Pro Glu Ser Asn Ser  
                   20                                  25                                  30

act cct cgt cct gcc gtc gct atc ggc gat ttg gtt ctg gac ctg tcc 144  
 Thr Pro Arg Pro Ala Val Ala Ile Gly Asp Leu Val Leu Asp Leu Ser  
                   35                                  40                                  45

gct atc tct gaa gct ggg ctt ttc gat ggt ctg atc ctt aag gac gca 192  
 Ala Ile Ser Glu Ala Gly Leu Phe Asp Gly Leu Ile Leu Lys Asp Ala  
   50                                  55                                  60

gat tgc ttt ctt cag cct aat ttg aat aag ttc ttg gcc atg gga cgg 240  
 Asp Cys Phe Leu Gln Pro Asn Leu Asn Lys Phe Leu Ala Met Gly Arg  
   65                                  70                                  75                                  80

## 125

cct gcg tgg aag gaa gcg cgt tct acg ctg caa aga atc ttg tca ttt	288
Pro Ala Trp Lys Glu Ala Arg Ser Thr Leu Gln Arg Ile Leu Ser Phe	
85 90 95	
ttg tta ttt ggc ttc aag gtt ttg gtt ttg gta tgt ttt cat gca gct	336
Leu Leu Phe Gly Phe Lys Val Leu Val Leu Val Cys Phe His Ala Ala	
100 105 110	
aat gaa cct atc ttg cga gac aat gat gtt ttg agg aga aaa tca ttc	384
Asn Glu Pro Ile Leu Arg Asp Asn Asp Val Leu Arg Arg Lys Ser Phe	
115 120 125	
cat cag atg agt aaa gtg gaa atg att gtt cct atg gtg att ggg gac	432
His Gln Met Ser Lys Val Glu Met Ile Val Pro Met Val Ile Gly Asp	
130 135 140	
tat aca gac ttc ttt gca tct atg cat cac gcg aag aac tgc gga ctt	480
Tyr Thr Asp Phe Phe Ala Ser Met His His Ala Lys Asn Cys Gly Leu	
145 150 155 160	
atg ttc cgt ggg cct gag aat gcg ata aac cca aat tgg ttt cgt ctt	528
Met Phe Arg Gly Pro Glu Asn Ala Ile Asn Pro Asn Trp Phe Arg Leu	
165 170 175	
ccc att gca tat cat gga cgg gca tca tct att gtc atc tct ggg act	576
Pro Ile Ala Tyr His Gly Arg Ala Ser Ser Ile Val Ile Ser Gly Thr	
180 185 190	
gac att att cga cca aga ggt cag ggc cat cca caa gga aac tct gaa	624
Asp Ile Ile Arg Pro Arg Gly Gln Gly His Pro Gln Gly Asn Ser Glu	
195 200 205	
cca tat ttt gga cct tcg aag aaa ctt gat ttt gag ctt gag atg gct	672
Pro Tyr Phe Gly Pro Ser Lys Lys Leu Asp Phe Glu Leu Glu Met Ala	
210 215 220	
gct gtg gtt ggt cca gga aat gaa ttg gga aag cct att gac gtg aat	720
Ala Val Val Gly Pro Gly Asn Glu Leu Gly Lys Pro Ile Asp Val Asn	
225 230 235 240	
aat gca gcc gat cat ata ttt ggt cta tta ctg atg aat gac tgg agt	768
Asn Ala Ala Asp His Ile Phe Gly Leu Leu Leu Met Asn Asp Trp Ser	
245 250 255	
gct agg gat att cag gcg tgg gag tat gta cct ctt ggt cct ttc ctg	816
Ala Arg Asp Ile Gln Ala Trp Glu Tyr Val Pro Leu Gly Pro Phe Leu	
260 265 270	
ggg aag agt ttt ggg act act ata tcc cct tgg att gtt acc ttg gat	864
Gly Lys Ser Phe Gly Thr Thr Ile Ser Pro Trp Ile Val Thr Leu Asp	
275 280 285	
gcg ctt gag cct ttt ggt tgt caa gct ccc aag cag gat cca cct cca	912
Ala Leu Glu Pro Phe Gly Cys Gln Ala Pro Lys Gln Asp Pro Pro Pro	
290 295 300	
ttg cca tat ttg gct gag aaa gag tct gta aat tac gat atc tcc ttg	960
Leu Pro Tyr Leu Ala Glu Lys Glu Ser Val Asn Tyr Asp Ile Ser Leu	
305 310 315 320	
gag cta gca cac cat acc gtt aac ggt tgc aat ttg agg cct ggt gat	1008
Glu Leu Ala His His Thr Val Asn Gly Cys Asn Leu Arg Pro Gly Asp	
325 330 335	
ctc ctt ggc aca gga acc ata agc gga ccg gag cca gat tca tat ggg	1056
Leu Leu Gly Thr Gly Thr Ile Ser Gly Pro Glu Pro Asp Ser Tyr Gly	
340 345 350	

## 126

tgc cta ctt gag ttg aca tgg aat gga cag aaa cct cta tca ctc aat 1104  
 Cys Leu Leu Glu Leu Thr Trp Asn Gly Gln Lys Pro Leu Ser Leu Asn  
 355 360 365

gga aca act cag acg ttt ctc gaa gac gga gac caa gtc acc ttc tca 1152  
 Gly Thr Thr Gln Thr Phe Leu Glu Asp Gly Asp Gln Val Thr Phe Ser  
 370 375 380

ggt gta tgc aag gga gat ggt tac aat gtt ggg ttt gga aca tgc aca 1200  
 Gly Val Cys Lys Gly Asp Gly Tyr Asn Val Gly Phe Gly Thr Cys Thr  
 385 390 395 400

ggg aaa att gtt cct tca ccg cct tga 1227  
 Gly Lys Ile Val Pro Ser Pro Pro  
 405

&lt;210&gt; 121

&lt;211&gt; 408

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis thaliana

&lt;400&gt; 121

Met Ala Leu Leu Lys Ser Phe Ile Asp Val Gly Ser Asp Ser His Phe  
 1 5 10 15

Pro Ile Gln Asn Leu Pro Tyr Gly Val Phe Lys Pro Glu Ser Asn Ser  
 20 25 30

Thr Pro Arg Pro Ala Val Ala Ile Gly Asp Leu Val Leu Asp Leu Ser  
 35 40 45

Ala Ile Ser Glu Ala Gly Leu Phe Asp Gly Leu Ile Leu Lys Asp Ala  
 50 55 60

Asp Cys Phe Leu Gln Pro Asn Leu Asn Lys Phe Leu Ala Met Gly Arg  
 65 70 75 80

Pro Ala Trp Lys Glu Ala Arg Ser Thr Leu Gln Arg Ile Leu Ser Phe  
 85 90 95

Leu Leu Phe Gly Phe Lys Val Leu Val Leu Val Cys Phe His Ala Ala  
 100 105 110

Asn Glu Pro Ile Leu Arg Asp Asn Asp Val Leu Arg Arg Lys Ser Phe  
 115 120 125

His Gln Met Ser Lys Val Glu Met Ile Val Pro Met Val Ile Gly Asp  
 130 135 140

Tyr Thr Asp Phe Phe Ala Ser Met His His Ala Lys Asn Cys Gly Leu  
 145 150 155 160

Met Phe Arg Gly Pro Glu Asn Ala Ile Asn Pro Asn Trp Phe Arg Leu  
 165 170 175

Pro Ile Ala Tyr His Gly Arg Ala Ser Ser Ile Val Ile Ser Gly Thr  
 180 185 190

Asp Ile Ile Arg Pro Arg Gly Gln Gly His Pro Gln Gly Asn Ser Glu  
 195 200 205

Pro Tyr Phe Gly Pro Ser Lys Lys Leu Asp Phe Glu Leu Glu Met Ala  
 210 215 220

Ala Val Val Gly Pro Gly Asn Glu Leu Gly Lys Pro Ile Asp Val Asn  
 225 230 235 240

Asn Ala Ala Asp His Ile Phe Gly Leu Leu Leu Met Asn Asp Trp Ser  
 245 250 255

## 127

Ala Arg Asp Ile Gln Ala Trp Glu Tyr Val Pro Leu Gly Pro Phe Leu  
 260 265 270

Gly Lys Ser Phe Gly Thr Thr Ile Ser Pro Trp Ile Val Thr Leu Asp  
 275 280 285

Ala Leu Glu Pro Phe Gly Cys Gln Ala Pro Lys Gln Asp Pro Pro Pro  
 290 295 300

Leu Pro Tyr Leu Ala Glu Lys Glu Ser Val Asn Tyr Asp Ile Ser Leu  
 305 310 315 320

Glu Leu Ala His His Thr Val Asn Gly Cys Asn Leu Arg Pro Gly Asp  
 325 330 335

Leu Leu Gly Thr Gly Thr Ile Ser Gly Pro Glu Pro Asp Ser Tyr Gly  
 340 345 350

Cys Leu Leu Glu Leu Thr Trp Asn Gly Gln Lys Pro Leu Ser Leu Asn  
 355 360 365

Gly Thr Thr Gln Thr Phe Leu Glu Asp Gly Asp Gln Val Thr Phe Ser  
 370 375 380

Gly Val Cys Lys Gly Asp Gly Tyr Asn Val Gly Phe Gly Thr Cys Thr  
 385 390 395 400

Gly Lys Ile Val Pro Ser Pro Pro  
 405

&lt;210&gt; 122

&lt;211&gt; 11667

&lt;212&gt; DNA

&lt;213&gt; Künstliche Sequenz

&lt;220&gt;

<223> Beschreibung der künstlichen Sequenz: supression  
 construct 2 p3300.1-Toc159-GFP-RNAi

&lt;400&gt; 122

aattcgtttc tccataataa tgtgtgagta gttcccagat aagggaatta gggttcctat 60  
 agggtttcgc tcatgtgttg agcatataag aaacccttag tatgtatttg tatttgtaaa 120  
 atacttctat caataaaatt tctaattcct aaaacaaaaa tccagtacta aaatccagat 180  
 cccccgaatt aattcggcgt taattcagca attcgtaatc atgggtcatag ctgtttcctg 240  
 tgtgaaattg ttatccgctc acaattccac acaacatacg agccggaagc ataaagtgtg 300  
 aagcctgggg tgcctaataa gtgagctaac tcacattaat tgcgttgccg tcaactgccc 360  
 ctctccagtc gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga 420  
 gaggcgggtt gcgtattggc tagagcagct tgccaacatg gtggagcacg acactctcgt 480  
 ctactccaag aatatcaaag atacagtctc agaagaccaa agggctattg agacttttca 540  
 acaaagggtg atatcgggaa acctcctcgg attccattgc ccagctatct gtcacttcat 600  
 caaaaggaca gtagaaaagg aagggtggc acacaaatgc catcattgag ataaaggaaa 660  
 ggctatcggt caagatgcct ctgccgacag tgggtcccaa gatggacccc caccacagag 720  
 gagcatcggt gaaaaagaag acgttccaac cacgtcttca aagcaagtgg attgatgtga 780  
 taacatgggt gagcacgaca ctctcgtcta ctccaagaat atcaaagata cagtctcaga 840  
 agaccaaagg gctattgaga cttttcaaca aagggttaata tcgggaaacc tcctcggatt 900  
 ccattgccc a gctatctgtc acttcatcaa aaggacagta gaaaaggaag gtggcaccta 960  
 caaatgccat cattgcgata aaggaaaggc tatcgttcaa gatgcctctg ccgacagtgg 1020  
 tcccaaagat ggacccccac ccacgaggag catcgtggaa aaagaagacg ttccaaccac 1080  
 gtcttcaaag caagtggatt gatgtgatat ctccactgac gtaagggatg acgcacaatc 1140  
 ccactatcct tcgcaagacc ttctctata taaggaagtt catttcattt ggagaggaca 1200  
 cgctgaaatc accagtctct ctctacaaat ctatctctct cgagtctacc atgagcccag 1260  
 aacgacgcc ggccgacatc cgccgtgcca ccgaggcgga catgccggcg gtctgcacca 1320  
 tcgtcaacca ctacatcgag acaagcacgg tcaacttcgg taccgagccg cagggaaccg 1380

aggagtggac	ggacgacctc	gtccgtctgc	gggagcgcta	tccctggctc	gtcgccgagg	1440
tggacggcga	ggtcgccggc	atcgccctacg	cgggcccctg	gaaggcacgc	aacgcctacg	1500
actggacggc	cgagtcgacc	gtgtacgtct	ccccccgcca	ccagcggacg	ggactgggct	1560
ccacgctcta	cacccacctg	ctgaagtccc	tggaggcaca	gggcttcaag	agcgtggctc	1620
ctgtcatcgg	gctgcccac	gacccgagcg	tgcgcacgca	cgaggcgctc	ggatatgccc	1680
cccgcggcat	gctgcgggcg	gccggcttca	agcacgggaa	ctggcatgac	gtgggtttct	1740
ggcagctgga	cttcagcctg	ccggtaccgc	cccgctccgt	cctgcccgtc	accgagattt	1800
gactcgagtt	tctccataat	aatgtgtgag	tagttccag	ataaggaat	tagggttcct	1860
atagggtttc	gctcatgtgt	tgagcatata	agaaaccctt	agtatgtatt	tgtatttgta	1920
aaatacttct	atcaataaaa	tttctaattc	ctaaaaccaa	aatccagtac	taaaatccag	1980
atcccccgaa	ttaattcggc	gttaattcag	tacattaaaa	acgtccgcaa	tgtgttatta	2040
agttgtctaa	gcgtcaattt	gtttacacca	caatatatcc	tgccaccagc	cagccaacag	2100
ctccccgacc	ggcagctcgg	cacaaaatca	ccactcgata	caggcagccc	atcagtcagg	2160
gacggcgta	gcgggagagc	cgttgtaagg	cggcagactt	tgctcatggt	accgatgcta	2220
ttcggaagaa	cggcaactaa	gctgccgggt	ttgaaacacg	gatgatctcg	cggagggtag	2280
catgttgatt	gtaacgatga	cagagcgttg	ctgcctgtga	tcaccgcggg	ttcaaaatcg	2340
gctccgtcga	tactatgtta	tacgccaact	ttgaaaacaa	ctttgaaaaa	gctgttttct	2400
ggtattttaag	gttttagaat	gcaaggaaca	gtgaattgga	gttcgtcttg	ttataattag	2460
cttcttgggg	tatctttaaa	tactgtagaa	aagaggaagg	aaataataaa	tggctaaaaa	2520
gagaatatca	ccggaattga	aaaaactgat	cgaaaaatac	cgctgcgtaa	aagatacggg	2580
aggaatgtct	cctgctaagg	tatataagct	ggtgggagaa	aatgaaaacc	tatatttaaa	2640
aatgacggac	agccggtata	aagggaccac	ctatgatgtg	gaacgggaaa	aggacatgat	2700
gctatggctg	gaaggaaagc	tgctgttcc	aaaggtcctg	cactttgaac	ggcatgatgg	2760
ctggagcaat	ctgctcatga	gtgaggccga	tggcgctcct	tgctcggaag	agtatgaaga	2820
tgaacaaagc	cctgaaaaga	ttatcgagct	gtatgocggag	tgcatcaggc	tctttcactc	2880
catcgacata	tccgattgtc	cctatacgaa	tagcttagac	agccgcttag	ccgaattgga	2940
ttacttactg	aataacgac	tggccgatgt	ggattgcgaa	aactgggaag	aagacactcc	3000
atttaaagat	ccgcgcgagc	tgtatgattt	tttaaagacg	gaaaagcccc	aagaggaact	3060
tgtcttttcc	cacggcgacc	tgggagacag	caacatcttt	gtgaaagatg	gcaaagtaag	3120
tggcttttatt	gatcttggga	gaagcggcag	ggcgacaaag	tggatatgaca	ttgccttctg	3180
cgtccggtcg	atcagggagg	atatccggga	agaacagtat	gtcgagctat	tttttgactt	3240
actggggatc	aagcctgatt	gggagaaaaa	aaaatattat	attttactgg	atgaattggt	3300
ttagtaccta	gaatgcatga	ccaaaatccc	ttaacgtgag	ttttcgttcc	actgagcgtc	3360
agaccccgta	gaaaagatca	aaggatcttc	ttgagatcct	ttttttctgc	gcgtaatctg	3420
ctgcttgcaa	acaaaaaaac	caccgctacc	agcggtggtt	tgtttgccgg	atcaagagct	3480
accaactctt	tttccgaagg	taactggctt	cagcagagcg	cagataccaa	atactgtcct	3540
tctagtgtag	ccgtagttag	gccaccactt	caagaactct	gtagcaccgc	ctacatacct	3600
cgctctgcta	atcctgttac	cagtggctgc	tgccagtggc	gataagtcgt	gtcttaccgg	3660
gttggaactca	agacgatagt	taccggataa	ggcgcagcgg	tcgggctgaa	cgggggggtt	3720
gtgcacacag	cccagcttgg	agcgaacgac	ctacaccgaa	ctgagatacc	tacagcgtga	3780
gctatgagaa	agcgccacgc	ttcccgaagg	gagaaaggcg	gacaggtatc	cggtaagcgg	3840
cagggctcga	acaggagagc	gcacgaggga	gcttccaggg	ggaaacgcct	ggtatcttta	3900
tagtctgtgc	gggtttcgcc	acctctgact	tgagcgtcga	tttttgtgat	gctcgtcagg	3960
ggggcgggagc	ctatggaaaa	acgccagcaa	cgcggccttt	ttacggttcc	tggccttttg	4020
ctggcctttt	gctcacatgt	tctttcctgc	gttatccctt	gattctgtgg	ataaccgtat	4080
taccgccttt	gagtgtgctg	ataccgctcg	ccgcagccga	acgaccgagc	gcagcagctc	4140
agtgagcgag	gaagcggaag	agcgctctgat	gcggtatctt	ctccttacgc	atctgtgcgg	4200
tatttcacac	cgcatatggt	gcactctcag	tacaatctgc	tctgatgccg	catagttaag	4260
ccagtataca	ctccgctatc	gctacgtgac	tgggtcatgg	ctgcgccccg	acacccgcca	4320
acacccgctg	acgcgccctg	acgggcttgt	ctgctcccgg	catccgctta	cagacaagct	4380
gtgaccgtct	ccgggagctg	catgtgtcag	aggttttcac	cgtcatcacc	gaaacgcgcg	4440
aggcaggggtg	ccttgatgtg	ggcgccggcg	gtcgagtggc	gacggcgcg	cttgtccgcg	4500
ccctggtaga	ttgcctggcc	gtaggccagc	cattttttgag	cggccagcgg	ccgcgatagg	4560
ccgacgcgaa	gcggcggggc	gtagggagcg	cagcgaccga	agggtagggc	ctttttgcag	4620
ctcttcggct	gtgcgctggc	cagacagtta	tgcacaggcc	agggcggttt	taagagtttt	4680
aataagtttt	aaagagtttt	aggcggaaaa	atcgcccttt	ttctctttta	tatcagtcac	4740
ttacatgtgt	gaccggttcc	caatgtacgg	ctttgggttc	ccaatgtacg	ggttccgggt	4800

cccaatgtac	ggcttttgggt	tcccaatgta	cgtgctatcc	acaggaaaga	gacctttttcg	4860
accttttttcc	cctgctaggg	caatttgccc	tagcatctgc	tccgtacatt	aggaaccggc	4920
ggatgcttcg	ccctcgatca	ggttgccgta	gcgcatgact	aggatcgggc	cagcctgccc	4980
cgcctcctcc	ttcaaategt	actccggcag	gtcatttgac	ccgatcagct	tgcgcacagg	5040
gaaacagaac	ttcttgaact	ctccggcgct	gccactgctg	tcgtagatcg	tcttgaacaa	5100
ccatctggct	tctgccttgc	ctgcggcgcg	gcgtgccagg	cggtagagaa	aacggccgat	5160
gccgggatcg	atcaaaaagt	aatcgggggtg	aaccgtcagc	acgtccgggt	tcttgccctc	5220
tgtgatctcg	cggtagatcc	aatcagctag	ctcgatctcg	atgtactccg	gccgcccggg	5280
ttcgctcttt	acgatcttgt	agcggcta	caaggcttca	ccctcggata	ccgtcaccag	5340
gcggccgttc	ttggccttct	tcgtacgctg	catggcaacg	tgctggtgtg	ttaaccgaat	5400
gcaggtttct	accaggctcg	ctttctgctt	tccgccatcg	gtcgcgcggc	agaacttgag	5460
tacgtccgca	acgtgtggac	ggaacacgcg	gccgggcttg	tctcccttcc	cttcccggta	5520
tcggttcatg	gattcggtta	gatgggaaac	cgccatcagt	accaggctcg	aatcccacac	5580
actggccatg	ccggccggcc	ctgcggaaac	ctctacgtgc	ccgtctggaa	gctcgtagcg	5640
gatcacctcg	ccagctcgtc	ggtcacgctt	cgacagacgg	aaaacggcca	cgtccatgat	5700
gctgcgacta	tcgcgggtgc	ccacgtcata	gagcatcgga	acgaaaaaat	ctggttgctc	5760
gtcgcccttg	ggcggcttcc	taatcgacgg	cgacccgctg	gccggcggtt	gccgggattc	5820
tttgcggatt	cgatcagcgg	ccgcttgcca	cgattcacgg	ggcgctgctt	ctgcctcgat	5880
gctgtgccgc	tgggcgccct	gcgcggccct	caacttctcc	accaggctat	caccagcg	5940
cgcgcgatt	tgtaccgggc	cggatggttt	gcgaccgtca	cgccgattcc	tcgggcttgg	6000
gggttccagt	gccattgcag	ggccggcaga	caaccacg	gcttacgcct	ggccaaccgc	6060
ccgttccctcc	acacatgggg	cattccacgg	cgctcggtgc	tggttggttct	tgattttcca	6120
tgccgcctcc	tttagccgct	aaaattcatc	tactcattta	ttcatttgct	catttactct	6180
ggtagctg	cgatgtattc	agatagcagc	tcggtaatgg	tcttgccctg	gcgtaccgcg	6240
tacatcttca	gcttggtgtg	atcctccgcc	ggcaactgaa	agttgaccgc	cttcatggct	6300
ggcgtgtctg	ccaggctggc	caacgttgca	gccttgctgc	tgctgctgct	cggacggccg	6360
gcacttagcg	tggttggtgct	tttgcctcatt	ttctctttac	ctcatttaact	caaatagagt	6420
ttgattta	ttcagcggcc	agcgccctgga	cctcgcgggc	agcgtcgccc	tcgggttctg	6480
attcaagaac	ggttggtgccc	gcggcgccag	tgccctgggtg	gctcacgcgc	tgctgatac	6540
gggactcaag	aatgggcagc	tcgtaccggg	ccagcgcctc	ggcaacctca	ccgcgatgc	6600
gcgtgccttt	gatcgcccgc	gacacgacaa	aggccgcttg	tagccttcca	tcctgacct	6660
caatgcgctg	cttaaccagc	tccaccaggt	cggcggtggc	ccatatgtcg	taagggtctg	6720
gctgcaccgg	aatcagcacg	aagtcggctg	ccttgatcgc	ggacacagcc	aagtcgcgcg	6780
cctggggcgc	tcctgctgatc	actacgaagt	cgccgcggcc	gatggccttc	acgtcgcggg	6840
caatcgctcg	gcggctcgatg	ccgacaacgg	ttagcgggtg	atcttcccgc	acggccgcgc	6900
aatcgcgggc	actgccctgg	ggatcggaat	cgactaacag	aacatcgggc	ccggcgagtt	6960
gcaggggcgc	ggctagatgg	gttgcgatgg	tcgtcttgcc	tgaccgcgct	ttctggttaa	7020
gtacagcgat	aaccttcattg	cgcttcccctt	gcgtatttgt	ttatttactc	atcgcatcat	7080
atacgcacg	accgcatgac	gcaagctgtt	ttactcaa	acacatcacc	tttttagacg	7140
gcggcgctcg	gtttcttccag	cggccaagct	ggccggccag	gccgccagct	tgccatcaga	7200
caaaccggcc	aggatttcat	gcagccgcac	ggttgagacg	tgccgcggcg	gctcgaacac	7260
gtaccgggcc	gcgatcatct	ccgcctcgat	ctcttcggta	atgaaaaacg	gttcgtcctg	7320
gcgctcctgg	tgccggtttca	tgcttggttcc	tcttggcgtt	cattctcggc	ggccgcagg	7380
gcgtcgccct	cgggtcaatgc	gtcctcacgg	aaggcacgcg	gccgcctggc	ctcggtgggc	7440
gtcacttctt	cgctgcgctc	aagtgcgcgg	tacagggtcg	agcgatgcac	gccaaagcag	7500
gcagccgcct	ctttcacggg	gcggcccttc	tggtcgatca	gctcgcgggc	gtgcgcgac	7560
tgtgcggggg	tgagggtagg	gcgggggcca	aacttcacgc	ctcgggcctt	ggcggccctg	7620
cgcgcgctcc	gggtgcgggtc	gatgattagg	gaacgctcga	actcggcaat	gccggcgaa	7680
acgggtcaaca	ccatgcgggc	ggccggcggtg	gtggtgtcgg	cccacggctc	tgccaggcta	7740
cgcaggcccg	cgcgggcctc	ctggatgcgc	tcggcaatgt	ccagtaggtc	gtgggtgctg	7800
cgggcccagg	ggtctagcct	ggtcactgtc	acaacgtcgc	cagggcgtag	gtgggtcaagc	7860
atcctggcca	gtcccgggcg	gtcgcgcctg	gtgcgggtga	tcttctcgga	aaacagcttg	7920
gtgcagccgg	ccgcgtgcag	ttcgcccgct	tggttggtca	agtcctgggtc	gtcggtgctg	7980
acgcgggcat	agcccagcag	gccagcgggc	gcgctcttgt	tcattggcgta	atgtctccgg	8040
ttctagtcgc	aagtattcta	ctttatgcga	ctaaaacacg	cgacaagaaa	acgccaggaa	8100
aaggggcagg	cggcagcctg	tcgcgtaact	taggacttgt	gcgacatgtc	gttttcagaa	8160
gacggctgca	ctgaacgtca	gaagccgact	gcactatagc	agcggagggg	ttggatcaaa	8220



gtactttgat	cccgagggga	accctgtggt	tggcatgcac	atacaaattgg	acgaacggat	8280
aaaccttttc	acgccctttt	aaatatccgt	tattctaata	aacgctcttt	tctcttaggt	8340
ttacccgccca	atatatcctg	tcaaacactg	atagtttaaa	ctgaaggcgg	gaaacgacaa	8400
tctgatccaa	gctcaagctg	ctctagcatt	cgccattcag	gctgcgcaac	tgttgggaag	8460
ggcgatcggg	gcgggcctct	tcgctattac	gccagctggc	gaaaggggga	tgtgctgcaa	8520
ggcgattaag	ttgggtaacg	ccagggtttt	cccagtcacg	acgttgtaaa	acgacggcca	8580
gtgccaagct	tttgggctaga	gcagcttgcc	aacatgggtg	agcacgacac	tctcgtctac	8640
tccaagaata	tcaaagatac	agtctcagaa	gaccaaaggg	ctattgagac	ttttcaacaa	8700
agggtaatat	cgggaaacct	cctcggtatc	cattgcccag	ctatctgtca	cttcatcaaa	8760
aggacagtag	aaaaggaagg	tggcacctac	aatgcccac	attgcgataa	aggaaaggct	8820
atcggtcaag	atgcctctgc	cgacagtggg	cccaaagatg	gacccccacc	cacgaggagc	8880
atcggtgaaa	aagaagacgt	tccaaccacg	tcttcaaagc	aagtggattg	atgtgataac	8940
atggtggagc	acgacactct	cgtctactcc	aagaatatca	aagatacagt	ctcagaagac	9000
caaagggcta	ttgagacttt	tcaacaaagg	gtaatatcgg	gaaacctcct	cggattccat	9060
tgcccagcta	tctgtcactt	catcaaaaagg	acagtagaaa	aggaagggtg	cacctacaaa	9120
tgccatcatt	gcgataaagg	aaaggctatc	gttcaagatg	cctctgccga	cagtgggtccc	9180
aaagatggac	ccccacccac	gaggagcatc	gtggaaaaag	aagacggttc	aaccacgtct	9240
tcaaagcaag	tggattgatg	tgatatctcc	actgacgtaa	gggatgacgc	acaatcccac	9300
tatccttcgc	aagaccttcc	tctatataag	gaagttcatt	tcatttgtag	aggacacgct	9360
gaaatcacca	gtctctctct	acaaatctat	ctctccatgg	catgttctgc	aggctgactc	9420
tagaggatcc	ccgggtaccg	agctcgaaga	tcttcgacgt	cgggaattcat	ggactcaaa	9480
tcggttactc	cagaaccaac	caaccccttc	tacgcttctt	cggggcaatc	aggaaaaacc	9540
tatgcttctg	ttgtcgccgc	cgtctgtgct	gcagccgccc	ataaggagga	tgggtggtgct	9600
gtgagtagtg	ccaaggagtt	ggattcctca	tcggaggctg	tgtctggtaa	ttcgggataag	9660
ggtggagctg	atgatttatc	tgactccgag	aaggagaagc	cgaatttggt	gggtgatggg	9720
aaggtttccg	acgaggtgga	tggttcttta	aaggaggatt	ctactactcc	tgaggctact	9780
ccgaagcctg	agggtggttt	tggtgagaca	attggtgtag	atgatgtttc	atcgttatct	9840
ccgaagccgg	aggctgtttc	tgatggtgta	ggggttggtg	aggagaataa	gaagggttaag	9900
gaggacgtgg	aggatattaa	agacgatggg	gagagtaaga	ttgaaaatgg	gagtgttgat	9960
ggtgatgtga	aacaggcttc	cacagatggg	gagagtga	aagcttccaa	cacttgtcac	10020
tactttctct	tatggtgttc	aatgcttttc	aagataccca	gatcatatga	aacggcatga	10080
cttcttcaag	agcgccatgc	ctgagggata	cgtgcaggag	aggaccatct	tcttcaagga	10140
cgacgggaac	tacaagacac	gtgctgaagt	caagtttgag	ggagacaccc	tcgtcaacag	10200
gatcgagctt	aagggaatcg	atttcaagga	ggacggaaac	atcctcggcc	acaagttgga	10260
atacaactac	aactcccaca	acgtatacat	catggccgac	aagcaaaaga	acggcatcaa	10320
agccaacttc	aagacccgcc	acaacatcga	agacggcggc	gtgcaactcg	ctgatcatta	10380
tcaacaaaat	actccaattg	gcgatggccc	tgtcctttta	ccagacaacc	attacctgtc	10440
cacacaatct	gccctttcga	aagatcccac	cgaagagaga	gaccacatgg	tccttcttga	10500
gtttgttaaca	gctgctggga	ttacacatgg	catggatgaa	ctatacaaac	atgatgagct	10560
ttaaggatcc	ttaaagctca	tcatgtttgt	atagttcatc	catgccatgt	gtaatcccag	10620
cagctgtttac	aaactcaaga	aggaccatgt	ggtctctctt	ttcgggtggga	tctttcgaaa	10680
gggcagattg	tgtggacagg	taatggttgt	ctggtaaaag	gacagggcca	tcgccaattg	10740
gagtattttg	ttgataatga	tcagcgagtt	gcacgcggcc	gtcttcgatg	ttgtggcggg	10800
tcttgaagtt	ggctttgatg	ccgttctttt	gcttgctggc	catgatgtat	acgttgtggg	10860
agttgtagtt	gtattccaac	ttgtggccga	ggatgtttcc	gtcctccttg	aaatcgattc	10920
ccttaagctc	gatcctgttg	acgaggggtg	ctccctcaaa	cttgacttca	gcacgtgtct	10980
tgtagttccc	gtcgtccttg	aagaagatgg	tcctctcctg	cacgtatccc	tcaggcatgg	11040
cgctcttgaa	gaagtcatgc	cgtttcatat	gatctgggta	tcttgaaaag	cattgaacac	11100
cataagagaa	agtagtgaca	agtgttgga	gctttctcac	tctccccatc	tgtggaagcc	11160
tgtttcacat	caacatcaac	actcccattt	tcaatcttac	tctcaccatc	gtctttaata	11220
tcctccacgt	cctccttaac	cttcttattc	tcctccacaa	cccctacacc	atcagaaaata	11280
gcctccggct	tcggagataa	cgatgaaaca	tcactacac	caattgtctc	accagaaacc	11340
acctcaggct	tcggagtagc	ctcaggagta	gtagaatcct	cctttaaaga	accatccacc	11400
tcgtcggaaa	ccttcccac	acccacaaa	ttcggcttct	ccttctcgga	gtcagataaa	11460
tcatcagctc	caaccttatc	cgaattacca	gacacagcct	ccgatgagga	atccaactcc	11520
ttggcactac	tcacagcacc	accatcctcc	ttatcggcgg	ctgcagcagc	agcggcgccg	11580

## 131

acaacagaag cataggtttt tcttgattgc cccgaagaag cgtagaaggg gttggttggt 11640  
 tctggagtaa ccgactttga gtccatg 11667

<210> 123

<211> 36

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
 oligonucleotide primer

<400> 123

ctcgaggaat tcatggactc aaagtcgggtt actcca 36

<210> 124

<211> 40

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
 oligonucleotide primer

<400> 124

ggatccataa gcaagctttc tcactctccc catctgtgga 40

<210> 125

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
 oligonucleotide primer

<400> 125

aagcttccaa cacttggtcac tacttt 26

<210> 126

<211> 26

<212> DNA

<213> Künstliche Sequenz

<220>

<223> Beschreibung der künstlichen Sequenz:  
 oligonucleotide primer

<400> 126

ggatccctaa agctcatcat gtttgt 26

# INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/02735

**A. CLASSIFICATION OF SUBJECT MATTER**  
IPC 7 C12N15/11 C12N15/82

According to International Patent Classification (IPC) or to both national classification and IPC

**B. FIELDS SEARCHED**

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 93 23551 A (SEYMOUR GRAHAM BARRON ;TUCKER GREGORY ALAN (GB); GRIERSON DONALD ( ) 25 November 1993 (1993-11-25) Ansprüche, examples 1-8 ---	1-17
Y	FIRE A ET AL: "Potent and specific genetic interference by double-stranded RNA in <i>Caenorhabditis elegans</i> " NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, vol. 391, 19 February 1998 (1998-02-19), pages 806-811, XP002095876 ISSN: 0028-0836 cited in the application the whole document --- -/--	1-17

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents:

\*A\* document defining the general state of the art which is not considered to be of particular relevance

\*E\* earlier document but published on or after the international filing date

\*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

\*O\* document referring to an oral disclosure, use, exhibition or other means

\*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\* & \* document member of the same patent family

Date of the actual completion of the international search

19 August 2003

Date of mailing of the international search report

27/08/2003

Name and mailing address of the ISA  
European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Kalsner, I

## INTERNATIONAL SEARCH REPORT

International Application No

PCT/EP 03/02735

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 02 00894 A (CROPDESIGN N V ;ZHOU ZHONGYI (BE); BROEKAERT WILLEM (BE); MIRONOV) 3 January 2002 (2002-01-03) the whole document ---	1-17
A	FIRE A: "RNA-triggered gene silencing" TRENDS IN GENETICS, ELSEVIER SCIENCE PUBLISHERS B.V. AMSTERDAM, NL, vol. 15, no. 9, 1 September 1999 (1999-09-01), pages 358-363, XP004176656 ISSN: 0168-9525 the whole document ---	1-17
A	MONTGOMERY ET AL: "RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 95, December 1998 (1998-12), pages 15502-15507, XP002138441 ISSN: 0027-8424 the whole document ---	1-17
A	WESLEY S VARSHA ET AL: "Construct design for efficient, effective and high-throughput gene silencing in plants" PLANT JOURNAL, BLACKWELL SCIENTIFIC PUBLICATIONS, OXFORD, GB, vol. 27, no. 6, September 2001 (2001-09), pages 581-590, XP002187670 ISSN: 0960-7412 the whole document -----	1-17

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/EP 03/02735

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 9323551	A	25-11-1993	AU 4079493 A	13-12-1993
			EP 0644942 A1	29-03-1995
			WO 9323551 A1	25-11-1993
			US 5942657 A	24-08-1999
			ZA 9303361 A	23-09-1994
WO 0200894	A	03-01-2002	AU 9165601 A	08-01-2002
			WO 0200894 A2	03-01-2002

Form PCT/ISA/210 (patent family annex) (July 1992)

## INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 03/02735

A. KLASSIFIZIERUNG DES ANMELDUNGSGEGENSTANDES  
 IPK 7 C12N15/11 C12N15/82

Nach der Internationalen Patentklassifikation (IPK) oder nach der nationalen Klassifikation und der IPK

## B. RECHERCHIERTE GEBIETE

Recherchierter Mindestprüfstoff (Klassifikationssystem und Klassifikationssymbole)  
 IPK 7 C12N

Recherchierte aber nicht zum Mindestprüfstoff gehörende Veröffentlichungen, soweit diese unter die recherchierten Gebiete fallen

Während der internationalen Recherche konsultierte elektronische Datenbank (Name der Datenbank und evtl. verwendete Suchbegriffe)

EPO-Internal, WPI Data

## C. ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Bez. Anspruch Nr.
Y	WO 93 23551 A (SEYMOUR GRAHAM BARRON ;TUCKER GREGORY ALAN (GB); GRIERSON DONALD ( ) 25. November 1993 (1993-11-25) Ansprüche, Beispiele 1-8	1-17
Y	FIRE A ET AL: "Potent and specific genetic interference by double-stranded RNA in <i>Caenorhabditis elegans</i> " NATURE, MACMILLAN JOURNALS LTD. LONDON, GB, Bd. 391, 19. Februar 1998 (1998-02-19), Seiten 806-811, XP002095876 ISSN: 0028-0836 in der Anmeldung erwähnt das ganze Dokument	1-17

☒ Weitere Veröffentlichungen sind der Fortsetzung von Feld C zu entnehmen

☒ Siehe Anhang Patentfamilie

\* Besondere Kategorien von angegebenen Veröffentlichungen :

\*A\* Veröffentlichung, die den allgemeinen Stand der Technik definiert, aber nicht als besonders bedeutsam anzusehen ist

\*E\* älteres Dokument, das jedoch erst am oder nach dem internationalen Anmeldedatum veröffentlicht worden ist

\*L\* Veröffentlichung, die geeignet ist, einen Prioritätsanspruch zweifelhaft erscheinen zu lassen, oder durch die das Veröffentlichungsdatum einer anderen im Recherchenbericht genannten Veröffentlichung belegt werden soll oder die aus einem anderen besonderen Grund angegeben ist (wie ausgeführt)

\*O\* Veröffentlichung, die sich auf eine mündliche Offenbarung, eine Benutzung, eine Ausstellung oder andere Maßnahmen bezieht

\*P\* Veröffentlichung, die vor dem internationalen Anmeldedatum, aber nach dem beanspruchten Prioritätsdatum veröffentlicht worden ist

\*T\* Spätere Veröffentlichung, die nach dem internationalen Anmeldedatum oder dem Prioritätsdatum veröffentlicht worden ist und mit der Anmeldung nicht kollidiert, sondern nur zum Verständnis des der Erfindung zugrundeliegenden Prinzips oder der ihr zugrundeliegenden Theorie angegeben ist

\*X\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann allein aufgrund dieser Veröffentlichung nicht als neu oder auf erfinderischer Tätigkeit beruhend betrachtet werden

\*Y\* Veröffentlichung von besonderer Bedeutung; die beanspruchte Erfindung kann nicht als auf erfinderischer Tätigkeit beruhend betrachtet werden, wenn die Veröffentlichung mit einer oder mehreren anderen Veröffentlichungen dieser Kategorie in Verbindung gebracht wird und diese Verbindung für einen Fachmann naheliegend ist

\*Z\* Veröffentlichung, die Mitglied derselben Patentfamilie ist

Datum des Abschlusses der Internationalen Recherche

19. August 2003

Absendedatum des internationalen Recherchenberichts

27/08/2003

Name und Postanschrift der Internationalen Recherchenbehörde  
 Europäisches Patentamt, P.B. 5818 Patentlaan 2  
 NL - 2280 HV Rijswijk  
 Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
 Fax: (+31-70) 340-3016

Bevollmächtigter Bediensteter

Kalsner, I

## INTERNATIONALER RECHERCHENBERICHT

Internationales Aktenzeichen

PCT/EP 03/02735

## C.(Fortsetzung) ALS WESENTLICH ANGESEHENE UNTERLAGEN

Kategorie*	Bezeichnung der Veröffentlichung, soweit erforderlich unter Angabe der in Betracht kommenden Teile	Betr. Anspruch Nr.
A	WO 02 00894 A (CROPDESIGN N V ;ZHOU ZHONGYI (BE); BROEKAERT WILLEM (BE); MIRONOV) 3. Januar 2002 (2002-01-03) das ganze Dokument	1-17
A	FIRE A: "RNA-triggered gene silencing" TRENDS IN GENETICS, ELSEVIER SCIENCE PUBLISHERS B.V. AMSTERDAM, NL, Bd. 15, Nr. 9, 1. September 1999 (1999-09-01), Seiten 358-363, XP004176656 ISSN: 0168-9525 das ganze Dokument	1-17
A	MONTGOMERY ET AL: "RNA as a target of double-stranded RNA-mediated genetic interference in Caenorhabditis elegans" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, Bd. 95, Dezember 1998 (1998-12), Seiten 15502-15507, XP002138441 ISSN: 0027-8424 das ganze Dokument	1-17
A	WESLEY S VARSHA ET AL: "Construct design for efficient, effective and high-throughput gene silencing in plants" PLANT JOURNAL, BLACKWELL SCIENTIFIC PUBLICATIONS, OXFORD, GB, Bd. 27, Nr. 6, September 2001 (2001-09), Seiten 581-590, XP002187670 ISSN: 0960-7412 das ganze Dokument	1-17

Formblatt PCT/ISA/210 (Fortsetzung von Blatt 2) (Juli 1992)

# INTERNATIONALER RECHERCHENBERICHT

Angaben zu Veröffentlichungen, die zur selben Patentfamilie gehören

Internationales Aktenzeichen

PCT/EP 03/02735

Im Recherchenbericht angeführtes Patentdokument	Datum der Veröffentlichung	Mitglied(er) der Patentfamilie	Datum der Veröffentlichung
WO 9323551 A	25-11-1993	AU 4079493 A	13-12-1993
		EP 0644942 A1	29-03-1995
		WO 9323551 A1	25-11-1993
		US 5942657 A	24-08-1999
		ZA 9303361 A	23-09-1994
WO 0200894 A	03-01-2002	AU 9165601 A	08-01-2002
		WO 0200894 A2	03-01-2002